Structural-Functional Studies in English Grammar

Edited by Mike Hannay and Gerard J. Steen
Structural-Functional Studies in English Grammar
Studies in Language Companion Series (SLCS)

The SLCS series has been established as a companion series to Studies in Language, International Journal, sponsored by the Foundation “Foundations of Language”.

Series Editors

Werner Abraham
University of Vienna

Michael Noonan
University of Wisconsin, Milwaukee

Editorial Board

Joan Bybee
University of New Mexico

Christian Lehmann
University of Erfurt

Ulrike Claudi
University of Cologne

Robert Longacre
University of Texas, Arlington

Bernard Comrie
Max Planck Institute
For Evolutionary Anthropology, Leipzig

Brian MacWhinney
Carnegie-Mellon University

William Croft
University of New Mexico

Marianne Mithun
University of California, Santa Barbara

Östen Dahl
University of Stockholm

Edith Moravcsik
University of Wisconsin, Milwaukee

Gerrit Dimmendaal
University of Leiden

Masayoshi Shibatani
Rice University and Kobe University

Ekkehard König
Free University of Berlin

Russell Tomlin
University of Oregon

Volume 83
Structural-Functional Studies in English Grammar
In honour of Lachlan Mackenzie
Edited by Mike Hannay and Gerard J. Steen
Structural-Functional Studies in English Grammar
In honour of Lachlan Mackenzie

Edited by
Mike Hannay
Gerard J. Steen
Vrije Universiteit Amsterdam

John Benjamins Publishing Company
Amsterdam/Philadelphia
Table of contents

Introduction 1

I. Corpus-based studies

No doubt and related expressions: A functional account 9
Anne-Marie Simon-Vandenbergen

On certainly and zeker 35
Pieter Byloo, Richard Kastein and Jan Nuyts

Prenominal possessives in English: Function and use 59
Evelien Keizer

Ditransitive clauses in English with special reference to Lancashire dialect 83
Anna Siewierska and Willem Hollmann

'It was you that told me that, wasn’t it? ’It'-clefts revisited in discourse 103
María de los Ángeles Gómez-González

Another take on the notion Subject 141
Dik Bakker and Anna Siewierska

The modal auxiliaries of English, π-operators in Functional Grammar and “grounding” 159
Louis Goossens

The king is on huntunge: On the relation between progressive and absentive in Old and Early Modern English 175
Casper de Groot

II. The architecture of functional models

Mental context and the expression of terms within the English clause: An approach based on Functional Discourse Grammar 193
John H. Connolly

Adverbial conjunctions in Functional Discourse Grammar 209
Kees Hengeveld and Gerry Wanders
# Table of contents

*Tree tigers and tree elephants: A constructional account of English nominal compounds*
  *Matthew Anstey*
  227

*English constructions from a Dutch perspective: Where are the differences?*
  *Arie Verhagen*
  257

*Notes towards an incremental implementation of the Role and Reference Grammar semantics-to-syntax linking algorithm for English*
  *Christopher S. Butler*
  275

*Grammar, flow and procedural knowledge: Structure and function at the interface between grammar and discourse*
  *Peter Harder*
  309

*The non-linearity of speech production*
  *Michael Fortescue*
  337

*A speaker/hearer-based grammar: The case of possessives and compounds*
  *Theo A. J. M. Janssen*
  353

*Index*

389
Introduction

This collection of studies is presented in honour of J. Lachlan Mackenzie, who in 2004 left the Chair of Modern English Language at the Vrije Universiteit, Amsterdam to start a new professional life. During a period of over twenty-five years working in English language and linguistics in the Netherlands, Lachlan made a significant contribution to the development of Simon Dik’s Functional Grammar and to functionalist linguistics in general. In addition to his theoretical work, Lachlan applied functional linguistic insights to the teaching of writing, and in recent years has played an increasingly prominent role in furthering the study of English in Europe as a whole with his work for the European Society for the Study of English (ESSE). Lachlan continues to work in Functional Grammar and is currently cooperating with Kees Hengeveld (Hengeveld & Mackenzie forthc.) on the architectural details of Functional Discourse Grammar (FDG), which has grown out of Functional Grammar (see Mackenzie & Gómez-González 2004). FDG takes units of discourse as the basic unit of analysis and aims to account more adequately than its predecessor for the form of linguistic expressions in terms of their use in actual communication.

Although he does not shy away from detailed study of other languages, Lachlan's work has been predominantly related to English, and his publications have come to be characterized by two strong features. In the first place he has insisted on fine-grained functional linguistic analysis, and in recent publications has made more and more use of authentic naturally occurring data for such purposes (e.g. Mackenzie 2003, 2005). Secondly, he has always sought to propose formal accounts which are true to the principles of the chosen theoretical framework, and which reinforce that framework. In this respect, his recent work in FDG has made proposals for a grammatical apparatus some elements of which operate in real time as part of a language production model, thus enhancing psychological adequacy (e.g. Mackenzie 2000, 2004).

In keeping with these interests, this collection honours him with a number of studies in the lexico-grammar of English which focus on the one hand on close reading of language in context and on the other hand on current functional theoretical concerns. The various contributions represent distinct functionalist models of language, including Functional Grammar and Functional Discourse Grammar, Systemic-Functional Grammar, Role and Reference Grammar, Cognitive Grammar and Construction Grammar. Taken together, however, they typify current work being conducted from the grammatical perspective within what Butler has called the functionalist enterprise, the aim of which is
to understand how the forms of human linguistic communication are related to the functions they serve in the exchange of meanings under conditions defined by the social and cognitive contexts of use, and by the structure of the ongoing interaction itself. (Butler 2003, Vol. 2: 477)

The core of the functionalist enterprise is indeed the relation between structure and usage. Although usage-based approaches have been attacked, notably by Newmeyer (2003), as confusing the boundary between knowledge of language and use of language, it remains a fundamental part of the enterprise to identify linguistic structures which are constrained by specific features of use, or which actually encode specific features of use, as many of the contributions here show.

Whatever particular theoretical model one embraces, the functionalist enterprise requires that that model should aspire in equal measure to descriptive adequacy and to various kinds of explanatory adequacy. In terms of descriptive adequacy this volume follows the call from Butler (2003) to embrace the study of corpora by offering a set of studies which are based on authentic linguistic data produced in well-defined natural communicative settings. Such studies permit a fine-grained linguistic analysis which can lead to enhanced descriptions of usage. In particular, insight into the array of contextual factors that play a role in determining usage may provide a basis for a formal account which contributes to the discoursal adequacy of the relevant theoretical model.

Part I accordingly offers a set of descriptively oriented, data-driven studies, starting off with Anne-Marie Simon-Vandenbergen’s functional account of no doubt and related expressions on the basis of corpus studies of both synchronic and diachronic data. No doubt has developed as a result of grammaticalization as a discourse marker, and the analysis of the context of use provides further evidence for the need to study modal expressions in general as means for negotiating opinions. Pieter Byloo, Richard Kastein, and Jan Nuyts look at a similar phenomenon, presenting a comparison of the use of certainly and Dutch zeker, on the basis of one of the Harry Potter books and its Dutch translation, plus two small samples from the British National Corpus (BNC) and the Corpus of Spoken Dutch (CGN). The close analysis, which also has an historical dimension, allows them to detail the similarities and differences between the two adverbs, and they speculate about what it is allows markers of epistemic certainty to develop a wider range of different meanings than weaker epistemic expressions. Evelien Keizer uses material taken from the British component of the International Corpus of English (ICE-GB) to examine the factors which determine the choice of prenominal possessives even if post-nominal alternatives also appear acceptable, and proposes a unifying cognitive-pragmatic principle which might account for the use of one form rather than another. Maria de los Ángeles Gómez-González focuses on the features and discourse distribution of it-clefts, also based on data from the ICE-GB corpus. She shows how the different configurations of given and new information within clefts relate to specific topic continuity strategies and proposes explanations for the distribution of clefts over different text types. Anna Siewierska and Willem Hollman consider
the coding of theme and recipient constituents in ditransitive clauses using four different corpora containing data from Lancashire dialect. They establish three rather than just two ditransitive patterns and search for evidence to determine whether the different constituent orders relate to pragmatic status or whether a more fundamental structural difference might be involved.

This first group of five studies all use corpora to gain a fuller understanding of how specific structures are used in a wider discourse context. In the three remaining studies in Part I, the authors go a step further and make use of corpus data to propose a reanalysis of how specific constructions might be handled in a functional model. First, Dik Bakker and Anna Siewierska tackle the controversial notion of subjecthood and use English corpus data to analyse how pragmatic and semantic factors determine speakers’ choice of a particular constituent as clausal subject. They go on to consider other languages and propose a multifunctional framework for subject choice which can be embedded in a functional model. Louis Goossens then uses corpus data to examine how modal auxiliaries of English are used over time, and exploits this analysis to argue for a modification of their cognitive-grammatical account within the model proposed by Langacker (1991) as well as their FG account within the model proposed by Dik (1997). Finally, Casper de Groot analyses a number of Middle English text fragments to examine how the use of two periphrastic constructions may be related to the progressive construction in present-day English, and exploits this analysis to propose a new hypothesis that competes with alternative proposals in the literature.

Part II goes yet a step further and is by contrast much less data-driven and more model-driven than Part I. The change of focus in this part means that a range of different models are involved and that the individual contributions are more geared to questions relating to explanatory rather than descriptive adequacy. The first contribution in Part II is from John Connolly, who shows how it is possible, in a model of the natural language user which relates FDG to a contextual component, to specify the relationship between these two components so as to provide an explanation for certain aspects of the use of adpositional expressions in English. Also working within FDG, Kees Hengeveld and Gerry Wanders then argue for the distinction of two different kinds of adverbial conjunction, lexical and grammatical, and illustrate how the interpersonal and representational levels proposed within FDG are relevant for modelling these different conjunction types. Matthew Anstey looks at English nominal compounds from the perspective of the reader/listener, and seeks to develop a model which can be psychologically adequate. To do this, he argues, it is necessary to adopt and expand principles of Construction Grammar. Arie Verhagen illustrates the value of a constructional approach in explaining the fact that grammars consist of both regularities and idiosyncrasies. This comes strongly to the fore when one compares similar constructions in two relatively closely related languages like Dutch and English. Finally in this group, Chris Butler takes as his starting point the incremental nature of language production and puts forward proposals for how the model of Role and Reference Grammar might gain in psychological adequacy by having a set of semantics-to-syntax mapping rules which capture the incrementality of the production process.
Butler’s contribution functions as a bridge to the last set of papers in Part II, which take a broader perspective. The concern here is with the nature of functional models as a whole, and in particular with the matter of the extent to which a model that seeks to account for language in communication needs a grammatical component which reflects the process of language production. Lachlan Mackenzie’s notion of an incremental Functional Grammar surfaces here. First, Peter Harder compares product-oriented and production-oriented models of language and evaluates the claims of FDG in these terms. For him it is essential that grammar be viewed as a set of procedures which are applied in the process of language production, rather than as a mechanism for modelling that process. Michael Fortescue also engages with a description of the production of speech, in particular the question of its linear or non-linear nature, and aims to show how a model of such a verbal process may be improved by taking into account some of the insights from the philosopher of language Arnold Whitehead. Finally, Theo Janssen discusses a number of basic issues in what he calls a speaker/hearer-based grammar, which aims to describe the use of utterances and their elements as a structured set of procedures. His concern is to promote a process-oriented model of language which combines both production and comprehension, and he seeks to describe the meaning and use of utterances as enriched and specified by the situational context.

This presentation, by a close colleague of Lachlan’s in Amsterdam, provides a fitting conclusion to the collection as it offers a singular take on the kind of grammar which is compatible with the functional enterprise. Indeed, the editors hope that all the articles collected here will provide readers with a stimulus to engage with the descriptive methods and modelling requirements associated with the functionalist enterprise.

The editors would like to express their sincere thanks to the series editors and to a number of external reviewers who made valuable comments on individual contributions and in one case on an earlier version of the complete manuscript.

Mike Hannay & Gerard Steen
Vrije Universiteit, Amsterdam

References


PART I

Corpus-based studies
No doubt and related expressions

A functional account*

Anne-Marie Simon-Vandenbergen

Ghent University

The epistemic adverb no doubt refers literally to absence of doubt, and hence to complete certainty, but it is actually described in the literature as expressing less than full certitude and as more or less synonymous with ‘very probably’. This contribution uses corpus data to examine the meaning and pragmatic uses of no doubt and of the formally and semantically closely related expressions there is no doubt and I have no doubt in present-day English. An explanation for how certainty adverbs such as no doubt have actually come to express some doubt or uncertainty is sought by considering data from historical corpora. It is argued that grammaticalization is a factor in this meaning development. But from a theoretical point of view the relationship between ‘certainty’ and ‘uncertainty’ needs an explanation which goes beyond the description of individual lexical items. This contribution shows that the explanation for such a relationship must be sought in the rhetorical exploitation of epistemic expressions.

1. Introduction

The epistemic adverb no doubt raises interesting descriptive and theoretical questions. Although it literally refers to absence of doubt, and hence to complete certainty, it is actually described in the literature as expressing less than full certitude. For instance, Quirk et al. point out that both doubtless and no doubt imply some doubt and are “syn-

* I wish to thank Karin Aijmer, Louis Goossens, Göran Kjellmer, an anonymous referee and the editors of this volume for their pertinent comments on an earlier version of this paper. They are, of course, not responsible for remaining shortcomings. I am indebted to Peter Flynn for drawing my attention to the ironical or sarcastic overtones of some types of no doubt utterances, and to Philippe Verelst for his help with the interpretation of Old French saunz doute. Thanks are also due to Bernard De Clerck and to Tine Defour for their generous assistance with collecting the corpus data.
Anonymous with 'very probably'” (1985:623, note [b]). Bolinger (1989:128) writes that such adverbs as *no doubt*, *surely* and *doubtless* are “stereotyped for 'uncertainty' despite the literal meaning of the words”. The first aim in this article is to verify the validity of these claims on the basis of corpus data from present-day English. Since epistemic certainty is scalar it is predictable that neat distinctions between the meanings 'certainly' and 'very probably' will not (always) be easy to make. Furthermore, Quirk et al.’s use of the words “imply some doubt” and Bolinger’s phrase “stereotyped for doubt” raise the question of the status of the meaning 'uncertainty'. Is it a meaning that is inferred in some contexts or is it part of the coded meaning?

The description of the meaning and use of *no doubt* needs to refer to the formally and semantically closely related expressions *there is no doubt* and *I have no doubt*. Both of these also occur frequently in present-day English and the exact nature of the semantic and pragmatic relationship between the three forms as well as the possible reasons speakers may have for choosing one expression rather than another in particular contexts calls for investigation on an empirical basis.

The second question that arises is how we can explain that some certainty adverbs come to express 'some doubt' (Quirk et al. 1985) or 'uncertainty' (Bolinger 1989). Aijmer (2002a) argues that “[s]uch meaning bifurcation is not unusual” in the area of modality, because of the unstable nature of the latter. She explains the development of the meanings of doubt or uncertainty, especially in the adverb *surely*, as the result of grammaticalization. From a manner adverb *surely* developed into a sentence adverb and a discourse marker. As a discourse marker *surely* is typically used “when the speaker is uncertain and seeks confirmation” (Aijmer 2002a:109). The development of *no doubt* is of a different nature in that its origin is not a manner adverb but a nominal group. In the clausal structures *there is no doubt* and *I have no doubt* this nominal group is syntactically integrated and these structures may have been the sources of the adverb through ellipsis (see Kjellmer 1998 on a similar alternation between *there is no question* and *no question*). A diachronic excursion will examine to what extent we can find evidence or indications of the historical relationship between *no doubt* and the clausal structures. An attempt will be made to account for the so-called 'uncertainty' meaning of *no doubt* on the basis of the historical and synchronic data.

From a theoretical point of view the relationship between 'certainty' and 'uncertainty' needs an explanation which goes beyond the details of individual modal elements. Halliday’s by now often quoted paradox that “we only say we are certain when we are not” (1994:362) is meant to refer to the marked nature of statements containing certainty expressions. Traugott and Dasher (2002:162) interpret it as meaning that "marking declarative signals some doubt about the truth of that declarative". However, the question remains what is meant by "some doubt". I want to argue that an account of the relationship between certainty and uncertainty needs to reinterpret 'doubt' in rhetorical terms rather than in epistemic terms.

The empirical data for the synchronic study have been taken from three corpora of present-day British English, viz. the British National Corpus (BNC), the International Corpus of English – Great-Britain (ICE-GB), and the Freiburg Lancaster-Oslo-Bergen
Corpus (FLOB). The BNC is ideal for showing general tendencies which do not emerge from smaller corpora. On the other hand its sheer size prevents arriving at accurate frequencies of particular patterns unless they can be unambiguously defined as search items for an automatic count. Since this is not possible in many cases, including the patterns in focus in the present study, smaller corpora whose automatic search output can be manually verified offer more reliable frequencies. The ICE-GB and FLOB corpora have been used for showing the relative frequencies of particular patterns. In the synchronic study attention is paid to syntax (position in the clause), lexis (collocations), and pragmatic factors (contextual clues guiding the interpretation). The diachronic discussion is based on two historical corpora, the Helsinki Corpus and the Corpus of Early English Correspondence Sampler.

2. A synchronic perspective

2.1 No doubt

2.1.1 Frequencies

Table 1 gives the frequency of the expressions no doubt, there is no doubt and I have no doubt in the ICE-GB and FLOB corpora.

The general tendencies are the same in both corpora, showing a clear preference for no doubt in present-day English. Least frequent is I have no doubt. These differences in frequencies are taken up again in Section 4.

1. The British National Corpus (BNC) is a 100 million word collection of samples of written and spoken language (period 1960–1993) from a wide range of sources, designed to represent a cross-section of current British English. For details see http://escorp.unizh.ch. ICE-GB is the British component of the International Corpus of English. It consists of a million words of spoken and written English. The texts were collected between 1990 and 1993. For information see http://www.ucl.ac.uk/english-usage/ice-gb. The Freiburg-LOB corpus matches the older Lancaster-Oslo-Bergen (LOB) corpus (which consists of texts from the 1970s) with texts from the 1990s. Like ICE-GB, it consists of one million words and is thus comparable with ICE-GB in terms of its size as well as the date of its materials. For more information see the site of the ICAME-corpora: http://helmer.aksis.uib.no/icame/manuals.html

2. The Helsinki Corpus of English Texts, Diachronic part (HC), covers the period from c. 750 to c. 1700. It includes a large selection of texts from the Old, Middle and Early Modern English period. In total it consists of approximately 1.5 million words. The Corpus of Early English Correspondence (CEECS) covers the timespan from 1417 to 1681, and the size of the whole corpus of letters is 2.7 million words. For information on both the HC and the CEECS see the site http://helmer.aksis.uib.no/icame/manuals.html
Table 1. Relative frequency of *no doubt* and related expressions in ICE-GB and FLOB

<table>
<thead>
<tr>
<th>Expression</th>
<th>ICE-GB</th>
<th>FLOB</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>no doubt</em></td>
<td>48 (61%)</td>
<td>39 (71%)</td>
</tr>
<tr>
<td><em>there is/can be no doubt</em></td>
<td>23 (29%)</td>
<td>10 (18%)</td>
</tr>
<tr>
<td><em>I have no doubt</em></td>
<td>8 (10%)</td>
<td>6 (11%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>79 (100%)</td>
<td>55 (100%)</td>
</tr>
</tbody>
</table>

Table 2. Positions of *no doubt*

<table>
<thead>
<tr>
<th>Position</th>
<th>ICE-GB</th>
<th>FLOB</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial</td>
<td>21 (44%)</td>
<td>15 (38%)</td>
</tr>
<tr>
<td>medial</td>
<td>25 (52%)</td>
<td>18 (46%)</td>
</tr>
<tr>
<td>final</td>
<td>2 (4%)</td>
<td>5 (13%)</td>
</tr>
<tr>
<td>alone</td>
<td>−</td>
<td>1 (3%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>48 (100%)</td>
<td>39 (100%)</td>
</tr>
</tbody>
</table>

2.1.2 Positions

Halliday (1994:82) claims that adjuncts of mood (including those of modality) occur in three basic positions: initial position (thematic), medial (neutral) and final (afterthought). Table 2 shows the frequency of *no doubt* in initial, medial and final positions in the clause.

Again, both corpora point in the same direction: medial position is most frequent (though the difference between initial and medial is fairly small), initial position is the second option, while in final position and as a response marker *no doubt* is much less frequent. In general, I have followed Quirk et al. (1985:490ff.) for classifying instances in terms of position.3

Initial position is defined in Quirk et al. (1985:491) as that “preceding any other element”. In this article instances of *no doubt* occurring as the first element in the clause as well as those following the co-ordinating conjunctions *and* and *but* (which are structurally restricted to initial position) have been classified as initial. In initial position *no doubt* is not normally followed by a comma. In ICE-GB there is only one instance, in FLOB none at all. In a random sample of 100 instances of initial *no doubt* in the BNC, not one had a comma.4 Separation by a comma indicates a lesser degree of integration in the clause. Disjuncts and conjuncts, for instance, are also typically

---

3. It should be noted that in classifying instances as initial, medial and final, only formal criteria are used. Thus any occurrence of *no doubt* (and the same applies to *there is no doubt* and *I have no doubt*) somewhere within a clause, i.e. non-thematic and non-final, has been counted as medial. I am aware that a more thorough analysis would also have to take into account the scope of the expressions. In some cases it is, for instance, clear that a medi ally occurring *no doubt* does not have the whole clause in its scope but only the following element.

4. In the spoken data, punctuation has obviously been inserted by the transcriber, either following a policy of transcription or following intuition. If the latter, intonation contours may be
No doubt and related expressions

No doubt and related expressions separated from the rest of the clause by commas (see the examples in Quirk et al. 1985:615ff.), reflecting an independent tone unit in speech. Fraser (1990:388) points out that initial now, set apart by a pause in speech and followed by a comma in writing, is a discourse marker. On the other hand, initial surely is not normally followed by a comma (see Downing’s (2001) examples). The results of prosodic analyses are inconclusive so far with regard to the question whether discourse markers form separate tone units (cf. Aijmer 2002b:32–33). Also, it is quite possible that results will vary from one discourse marker to another. More work needs to be done in this area before we can correlate functions with prosodic status. A study of the intonation contours of utterances with initial no doubt could show to what extent the adverb is prosodically integrated in the utterance.

Medial position as defined by Quirk et al. (1985:491ff.) comes in different variants, which share the occurrence of the adverbial somewhere between the subject and the main verb. This includes cases where the adverbial comes between the subject and the finite verb as well as cases where it occurs between the finite operator and non-finite auxiliaries or the non-finite main verb. Apart from such ‘prototypical’ medial cases (illustrated in example (1) below) I have, however, also classified as medial those instances where no doubt occurred within a phrase (as in (3), (4) and (5) below) and instances where it follows another thematic adverbial (as in (2) below). Finally, in contrast with Quirk et al. (1985:490), I have not made use of the category ‘initial end’, which is the position after the verb phrase and before an obligatory element such as an object. In the present study such instances have been classified as medial. It appears from the data that no doubt is extremely versatile. Such versatility characterizes discourse particles, which “are placed with great precision at different places in the discourse” (Aijmer 2002b:2). In medial positions no doubt is typically put between commas to mark its parenthetical nature. The following instances illustrate the extreme positional flexibility of no doubt:

1. Part of the Heseltine strategy is, no doubt, based on the calculation that these debts can be collected (…). (BNC, A5K, 308; written/non-fiction/newspaper) within the verb phrase, after the finite operator
2. Usually, no doubt, the audience shared the unstated assumptions, but even if they did not, it would have been a solecism to answer the question. (BNC, ADW, 329; written/non-fiction): after a thematic adverb
3. Thompson (1980) has warned of the potential danger of a strong police autocracy, conscious, no doubt, of the slim line which exists between the democratic use of power and its subversion by a more centralized totalitarianism. (BNC, A0K, 130; written/non-fiction): within an adjective phrase

significant. In any case we have to allow for the role of the transcriber in the case of the spoken examples. I am grateful to the anonymous referee for making this point.
(4) Scores of chained skeletons lay in the aisles – rivals, *no doubt*, of the regime, potential competitors for power who had been left here naked over the centuries (...). (BNC, CJ], 1720; written/fiction): within a noun phrase

(5) Because, *no doubt*, of the different structure of the population in Scotland, the Wheatley Commission took a different view. (BNC, ED5, 779; written/non-fiction): within a complex preposition

When final, *no doubt* is typically preceded by a comma, indicating its loose connection with the clause structure. From a discourse functional point of view, final *no doubt* has the status of an afterthought. Its perfunctory character is emphasized by its very frequent occurrence in a syntactically incomplete sentence. Ellipsis in various forms is most characteristic of dialogue (Halliday 1994:337), and this is where final *no doubt* belongs. It should be noted, however, that ‘dialogue’ must be interpreted to include interior monologue. The following instances illustrate final *no doubt*:

(6) ‘And you naturally think you should read them’, said Sir George.  
‘I – we – should like to, very much. Of course’.  
‘So would that American, *no doubt*’. (BNC, APR, 2166; written/fiction)

(7) ‘Why was Rahmi arrested?’
Jean-Pierre shrugged. ‘Subversion, *no doubt*. Anyway, Raoul Clermont is running around town trying to find Ellis and somebody wants revenge’. (BNC, CM7, 1018; written/fiction)

(8) What had he come out with? The unfaced truth, *no doubt*. (BNC, FSP, 1462; written/fiction)

The independent status of *no doubt* is most manifest in responses:

(9) ‘Got your husband picked out already?’
‘Not yet. But I can assure you I’ll give it my full attention.’
‘*No doubt*’. His smile was hard. (BNC, JYD, 1327; written/fiction)

In sum, both medial and final positions mark *no doubt* as an adverb which is syntactically unintegrated in the clausal structure and which has a parenthetical status. In those positions it is separated by commas. Swan points out that sentence modifiers “can easily be seen as reduced sentences with separate and separating intonation pattern/punctuation” (1988:512). In initial position *no doubt* is, however, not normally marked off from the rest of the clause by a comma. Such marking in thematic position would give it extra focus, which it seems to resist. *No doubt* is also used independently, as a response marker.

2.1.3 Collocations

*No doubt* frequently co-occurs with the auxiliaries *will* and *would*. The BNC has 222 instances of *will*/*ll no doubt*, in which *will* has the meaning of prediction or assumption. *Would* conveys prediction in the past or speculation about a hypothetical situation. In
such contexts no doubt clearly expresses the speakers’ conviction that the state of affairs took or will take place but that they are not in a position to have absolute certainty, because they do not possess any evidence. Consider examples (10)–(13) as illustrations of this use:

(10) The reign of the ultra-model will, no doubt, continue for a time. The relentless bombardment of images and icons of affluence and perfection is not waning. (BNC, CGB, 796, written/non-fiction)

(11) ‘No doubt, Mr. Deputy Speaker, you will wonder how rail transport could relate to the Scottish bus passengers’ consultative committee (…).’ (BNC, G3H, 1005; written/non-fiction/Weekly Hansard)

In (11) no doubt indicates unwillingness to state unequivocally what another person thinks, because the speaker has no access to the person's thoughts and can only speculate. Examples (1), (2), (3), (5) above illustrate the same function of no doubt.

(12) At that time Fothergill had, as he explained in a letter to Bartram, 1744, little chance to pursue the practical study of plants as he would wish, but he did, now and then, take a walk to Peckham or Chelsea. In those gardens he would, no doubt, have noted new plants from America and how best to grow them in this country. (BNC, ALU, 680; written/non-fiction)

(13) Mair said: ‘If the answer to that question were known, no doubt the appointment would have been made by now’. (BNC, CT8, 718; written/non-fiction)

The BNC has only two instances of the collocation may no doubt, but they are interesting because they point to the meaning of possibility rather than certainty. Example (14) illustrates this:

(14) It has to be admitted that the sound quality on some of these recordings at times leaves something to be desired, and some listeners may no doubt be disturbed by the occasional fluctuations of pitch, as in the Lohengrin Prelude. But I do not believe that H. Ward Marston's transfers are surpassable, and am grateful that the surface noise of the original discs remains intact rather than being artificially eliminated. (BNC, BMC, 2371; written/non-fiction)

On the other hand, the whole sentence in (14) is concessive and followed by another one which starts with but. There seems to be a merger in (14) between the epistemic and the concessive uses of may. We must conclude that epistemic may does not usually collocate with no doubt.

The collocation no doubt... but, on the other hand, is frequent. The clause containing no doubt has a concessive function: the information is backgrounded and made subordinate to another claim which is foregrounded. The contrastive conjunction but flags that there is a counter-claim. The expectation that a but-clause will follow is very strong in contexts where the speaker/writer uses no doubt in a clause which expresses a view which does not contribute to the argumentation. It may be either a claim uttered...
by another voice or a counter-argument brought up by the speaker/writer. Consider (15) to illustrate this use:

(15) As the geological cycles repeated themselves for some three thousand million years, as the volcanoes exploded and spent themselves, life in the sea burgeoned into many forms; but the land still remained barren. Some marine algae, no doubt, managed to live on the edges of the seas, rimming the beaches and boulders with green, but they could not have been spread far beyond the splash zone, for they would have dried out and died. (BNC, EFR, 738; written/non-fiction)

In the following instance this backgrounding function of no doubt is made explicit:

(16) Oliver had assured her that Sir Thomas could not possibly have committed suicide through anguish at her rejection and she was now prepared to enjoy the rest of the week immensely, even though it did involve a murder investigation. It was a shocking thing, no doubt, but compared with her future, which suddenly seemed to hold out enticing prospects of a new and exciting life, it had retreated into the background. (BNC, H8A, 2961; written/fiction)

In example (16) it is clear that no doubt does not express probability but certainty: the suicide being referred to is felt to be ‘a shocking thing’. The function of no doubt is to concede the truth in order to posit the counter-argument in a context of dialogic argumentation. In such contexts it is closer to certainly or of course than to probably (see Simon-Vandenbergen & Aijmer 2003 on the concessive use of of course).

The difference between such examples as (11) and (16) shows that no doubt may convey different degrees of epistemic certainty. Its meaning is not stable across contexts.

2.1.4 Contextual clues

In many cases there are contextual indications that no doubt expresses less than full certainty. Such clues include expressions of doubt, as in (17) below, where the character considers different possibilities and explicitly contrasts one hypothesis with an alternative one:

(17) Once, no doubt, there would have been a bridge, since swept away by floods or war. Or perhaps it had never existed. (BNC, HTT, 1823; written/fiction)

The following instance makes explicit the nature of the epistemic qualification by reference to the source of knowledge, viz. inferencing. ‘Reasonable inference’ indeed characterizes the meaning of no doubt well:

(18) It is a reasonable inference that the running of a market was profitable; so, no doubt, was the issuing of authorizations. (BNC, HPT, 369; written/non-fiction)
The speaker’s inference is often based on the predictability of a state of affairs. This predictability follows from expectations based on past experience. The following examples make this meaning aspect of ‘predictability’ explicit:

(19) The term ‘picaresque’ is used pretty loosely here, no doubt, as it often is, to mean something like episodic and comically adventurous. (BNC, CKN, 378; written/non-fiction)

(20) He was an old man now, and Doyle pointed this out. ‘Mellowed, too, no doubt,’ added Cowley. ‘They all do. Their criminal pasts become the “good old days”, stories for the children (. . .).’ (BNC, CE5, 677; written/fiction)

This meaning of ‘high degree of predictability’ leads to mild ridicule or even to sarcasm in some contexts. This is especially true when no doubt is used alone (as a response) or in final position. Its function is then similar to the use of would with the meaning of ‘one expects that’, as in He would say that, wouldn’t he? Many elliptical structures with final no doubt have this negative connotation. This type of no doubt typically occurs in conversation, where it functions as a discourse marker inviting confirmation. Consider examples (21) and (22):

(21) ‘You’re treating the whole thing like a military exercise,’ said Heidi, laughing. ‘All except the last part,’ said Adam. ‘The last part?’ she queried. ‘Our celebration dinner.’ ‘At the Chelsea Kitchen again, no doubt.’ ‘Wrong,’ said Adam. ‘I’ve booked a table for two at eight o’clock at the Coq d’Or just off Piccadilly.’ (BNC, EDV, 2450; written/fiction)

(22) ‘Well, what’s been happening in the club?’ asked Chatterton. ‘Wine, women and song, no doubt.’ ‘Oh no, Cully,’ said Glastonbury. ‘It’s been very dull since you left.’ (BNC, HTG, 869; written/fiction)

Summing up, the syntactic positions that no doubt takes point to its development into a parenthetical marker. It is an epistemic expression which can be thrown into the clause at almost every point where the speaker thinks fit. Further, its positions mark it as non-emphatic. Its meaning of ‘high degree of certainty but not complete certainty’ can be specified as ‘probability assessment based on a reasonable inference’. This means it is used when the speaker has inner certainty but no evidence. However, the degree of conviction is context-dependent. No doubt seems to slide on the epistemic certainty-probability scale. In some contexts it is synonymous with probably, in others with certainly. Further, it has been shown that no doubt is used as a discourse marker inviting confirmation from the addressee and giving a confirmatory response to a previous statement. There is in this respect a close parallel with surely, which is usually treated as an evidential (see Downing 2001) and has discourse marker characteristics (see Aijmer 2002a).
2.2 There is no doubt

2.2.1 Syntactic patterning
A comparison of the semantics and pragmatics of no doubt with those of there is no doubt and I have no doubt can throw light on why it is that a certainty marker expresses 'uncertainty'. One factor to examine is whether the two full clausal expressions have the same semantic and pragmatic versatility on the epistemic certainty-probability scale. Another factor is the syntactic behaviour of the clausal expressions. The question that interests me here is whether there are indications that they, too, show parenthetical behaviour, and to what extent they still have clausal status or have become epistemic markers. The aim is to examine the similarities and the differences between them.

There are two basic patterns: one in which no doubt is followed by a prepositional phrase which is the postmodifier in the nominal group of which doubt is the head; and one in which there is no doubt is followed by a clause. The two patterns can also be combined, as in there is no doubt about it that... Since it is the clausal pattern which is structurally closely related to a clause with no doubt, it is this pattern which will be further discussed here. There are two types of clauses following there is no doubt, a wh-clause and a that-clause (the latter with or without deletion of the conjunction). Consider the following example:

(23) There is no doubt where the blame lies. (BNC, A69, 1343; written/non-fiction)

With a following wh-clause the expression has full propositional status, since a question cannot be epistemically qualified (see e.g. Nuyts 2001:58 for a discussion of apparent exceptions). An interesting example from FLOB is the following:

(24) He looks at her in a slightly pleading fashion. No doubt who's the lover and the loved one. (FLOB, K08/99, written/fiction)

This instance of no doubt is clearly an elliptical existential rather than an adverb, and it shows that grammaticalization from the full clausal type is certainly a possible development.

With a that-clause, the existential structure still has full propositional status but its meaning is close to that of an epistemic qualifier:

(25) There is no doubt that Hitler's anti-Semitism, perceived as it was chiefly in connection with legal discrimination against Jews, was acceptable to the millions of his admirers. (BNC, ADD, 937; written/non-fiction)

The conjunction can be deleted:

(26) I have personally been down to the suppliers several times and checked, and there is no doubt we will have delivery. (BNC, FRS, 1309; written/fiction)

In a random sample of 100 BNC instances of there is no doubt there was not one case in which the expression occurred anywhere but in initial position in the sentence. There
No doubt and related expressions

Table 3. Relative frequency of that-deletion

<table>
<thead>
<tr>
<th>Expression</th>
<th>that</th>
<th>Deleted that</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>there is no doubt</td>
<td>82 (82%)</td>
<td>6 (6%)</td>
<td>12 (12%)</td>
<td>100</td>
</tr>
<tr>
<td>there’s no doubt</td>
<td>11 (50%)</td>
<td>4 (18%)</td>
<td>7 (32%)</td>
<td>22</td>
</tr>
</tbody>
</table>

was, however, one instance of reversal in a random sample of there can be no doubt instances:

(27) That New Zealand needs the investment there can be no doubt. (BNC, AJU, 921; written/non-fiction:newspaper)

There are 20 instances of there’s no doubt in the whole BNC, none of which occurs in final or medial position. What does occur is there’s no doubt in a paratactic relationship with the clause on which it comments. In these cases the postmodifier contains an anaphoric pronoun:

(28) He’s made a packet from that factory, there’s no doubt about that. (BNC, CJX, 2729; written/fiction)

(29) I don’t think stress would cause this sort of pain but it’ll certainly make it feel worse, there’s no doubt about that. (BNC, G5T, 85; spoken/medical consultation)

There’s no doubt about it/that also occurs on its own:

(30) yes, oh yes, th th that’s right, there’s no doubt about it. (BNC, KC3, 307; spoken/informal conversation)

It appears from the data that the existential structure shows some positional variability, but mostly occurs in initial position. The pattern there is no doubt that... is by far the most frequent one. Table 3 shows the frequency of clause-initial instances with and without the subordinator that as well as the category of all other types of complements in a random sample of 100 instances of there is no doubt and the total of 22 instances of there’s no doubt in the BNC.

What does Table 3 reveal about the status of there is no doubt? Aijmer (1997) points out that I think is followed by the conjunction that in only 7% of the cases in the London-Lund corpus, thus confirming Thompson and Mulac’s findings (1991) for American English. Aijmer comments that when that is used, I think is more likely to express “an informative statement about the speaker’s beliefs”, than to have an expressive function (Aijmer 1997:10). She further makes a distinction between the ‘deliberative’ and ‘tentative’ meanings of I think, and remarks that “all examples where I think is followed by the that-complementizer have been regarded as deliberative since they carry some prominence” (1997:21). Cases of I think in medial and final position she classifies as tentative. If we draw a parallel between I think, which developed into a modal particle (Aijmer 1997; Kärkkäinen 2003), and there is no doubt, it would appear that there are no signs that there is no doubt is developing in the same way. Thompson
and Mulac (1991:313) also consider cases where I think is followed by that as least grammaticalized, while I think without the subordinating conjunction comes closer to being a pragmatic element, and medial or final I think are seen as most grammaticalized. There is another factor which points in the same direction, i.e. which indicates that there is no doubt tends to keep its full propositional and assertive meaning. That factor is lexical variation and collocation.

2.2.2 Variation and collocation
When the existential expression is expanded by postmodifiers of the head noun doubt it tends to be assertive and emphatic. Here are some examples:

(31) There is no doubt about it that this programme is an Irish Republican Army programme sponsored and inspired by Communism. (BNC, APP, 517; written/non-fiction)

(32) Success brings its own problems and there is no doubt in my mind whatever that the Hunt we saw in 1977 was no longer the Hunt of 1976. (BNC, CD9, 1334; written/non-fiction)

(33) There is no doubt at all that where there are reductions er, particularly at the edge of our statutory responsibilities, this inevitably, we know, generates more demands for support from central services (...). (BNC, J3P, 687; spoken/public/institutional)

Such emphatic modifiers give the expression additional prominence, in the same way that I think is given prominence in combination with certain lexical elements (see Simon-Vandenbergen 1998:300). At the same time the variation points to full propositional status.

Another type of variation is with the auxiliary can. The BNC has 206 instances of there can be no doubt. Example (34) illustrates this pattern:

(34) There can be no doubt now that one member, one vote ballots will become compulsory. (BNC, A2A, 56; written/non-fiction/newspaper)

The expression can further be combined with epistemic markers which are explicitly subjective, i.e. which indicate that the assessment is the speaker’s only:

(35) There’s no doubt so it seems to me uh that for a considerable time he may well have hoped that uh some at least of these witnesses from Gibraltar who gave evidence at the inquest uh might’ve been in the case of (...). (ICE-GB, S2A-063/68; spoken/legal presentations)

(36) I think there’s no doubt of that. (ICE-GB, S2A-040/5, spoken/unscripted speeches/inaugural lecture)

The epistemic expressions I think and it seems to me qualify in a subjective way the proposition there’s no doubt. In other words, these collocations are further evidence
of the propositional nature of the existential expression. In the following instance the propositional meaning is highlighted by the contrast in the following sentence:

(37) There is no doubt that major volcanic eruptions have affected the climate of the northern hemisphere in the past few hundred years (see ‘Do volcanoes affect the climate?’ New Scientist, vol. 93, p. 150). The doubt concerns the speed with which the effects are felt, and the magnitude of those effects. (BNC, B7J, 1526; written/non-fiction/magazine)

Both syntactic and collocational factors indicate that there is no doubt has full propositional meaning. This would indicate that it has its literal meaning of intersubjective certainty (in the sense of Nuyts 2001, i.e. certainty shared by others), since it literally means the non-existence of doubt. Whether this is the case needs to be examined in the contexts in which the expression occurs.

2.2.3 Contextual meanings
In most cases it is indeed the case that the certainty expressed by there is no doubt is shared, intersubjective. Example (37) above is a case in point. Other examples are (38) and (39):

(38) In the last few years a powerful controversy has arisen over whether the dinosaurs as a whole were cold-blooded, like all living reptiles (and there is no doubt that dinosaurs were reptiles), or warm-blooded, resembling mammals in this respect. (BNC, AMM, 1769; written/non-fiction)

(39) The actual mechanics of establishing rapport have been investigated very carefully and there is no doubt that it is a behavioural phenomenon. (BNC, B2F, 1427; written/non-fiction)

Such instances show that there is no doubt expresses both an intersubjective (rather than subjective) assessment and absolute certainty (rather than probability). However, it is also clear that speakers do not express absolute absence of doubt unless they have a special reason for doing so (Traugott 1989). There seem to be two main types of reasons for expressing certainty in discourse. One is to contrast the shared certainty of one proposition with the controversial nature of another proposition, as in (40) below:

(40) We may not yet be a vigilante state but there is no doubt that more people feel compelled to mete out their own justice rather than entrust it to the police and courts. (BNC, CBC, 13616; written/non-fiction/newspaper)

The second main reason is that not everyone accepts the truth of the proposition, as illustrated in (41) below:

(41) And the idea that postural or emotional changes could bring on purring also seems far-fetched. The purring animal is relaxed, and common sense demands that this is precisely the time when any blood turbulence would subside, not increase. Finally, the false vocal cord theory is supported by the
evidence of the simple action of placing your fingers gently on the throat of a purring cat – **there is no doubt, then, that** the purring sound is stemming from the region of the voice-box, and it is hard to see why any alternative theory should have been put forward. That it has, despite the fact that superficially it is so unlikely, should, however, keep us on our toes. (BNC, BMG, 93; written/non-fiction)

In this passage various alternative theories are compared and argued for or against. The expression **there is no doubt** is used with a proposition which is presented as empirically verifiable. However, it is also made explicit that there is divergence of opinion on theories explaining cat purring, and the argumentative nature of the text illustrates the rhetorical need for expression of certainty when discussing controversial issues. A heteroglossic account (in White's 2003 sense) explains the occurrence of certainty markers when the assessment is subjective rather than intersubjective, and hence when there is no absolute, shared certitude. In this account, which draws on the Bakhtinian view that all communication is dialogic, modality is seen as a set of semantic options which speakers select from in order to position themselves with respect to other voices holding similar or different viewpoints. Modal choices are just one set of options within what is called the system of 'engagement'. What is crucial is that 'epistemic' choices are seen as not only and often not primarily expressing the speaker's state of knowledge but also and mainly as rhetorical means for constructing "authorial personae" (White 2003:260). Modal expressions construe heteroglossic engagement, in the sense that they open up the dialogue, by recognizing the existence of alternative views.

The accompaniment of subjectivity markers such as *it seems to me* and *I think* as in (35) and (36) above further emphasizes the lack of full certitude. Lexical expansions such as *in my mind* as in (32) above emphasize the subjective nature of the certainty. Example (42) below illustrates that in combination with a 'proposal' (in Halliday's sense 1994:71) rather than with a proposition, **there is no doubt in my mind** conveys an unequivocally subjective assessment:

(42) **There is no doubt in my mind** that if the Pakistan bowlers have transgressed the laws they should be stopped. (BNC, CH7, 5339; written/non-fiction/newspaper)

**There can be no doubt** literally expresses greater certainty than **there is no doubt** because it not merely refers to the absence of doubt but excludes the possibility of any doubt arising. It seems to be used especially in controversial contexts, where the speaker/writer uses extra emphasis to persuade. As a result, it is sometimes used to convey one viewpoint, the speaker's, against an alternative viewpoint. In the following example the speaker clearly expresses her own opinion, not a shared certainty:

(43) **There can be no doubt whatever that** the policies which would be pursued by the party opposite would work not only to the detriment of my honourable friend's constituents – very serious though that would be – but to the detri-
No doubt and related expressions

Summing up, various factors indicate that there is no doubt has not, in contrast with I think for example, developed into a modal particle. It tends to be clause initial, followed by the coordinating conjunction that, it shows internal lexical variation. Its meaning is that of full certainty, which is also typically intersubjective. However, rhetorical factors prompting its use make it useful in argumentative contexts where the certainty assessment is not shared by everyone. It follows that there is no doubt can express subjective as well as intersubjective certainty. However, even when subjective the assessment is one of certainty, not of probability, as is the case with no doubt. On the other hand, some indication of lack of certitude is brought in by markers such as I think and it seems to be. There are also contexts in which it is clear that the speaker has no hard and fast evidence but is expressing a plausible hypothesis. The heteroglossic function of there is no doubt explains how the adverb no doubt could have developed into a marker of probability. The distinction between assertive there is no doubt and tentative no doubt parallels that between assertive I think and tentative I think. In the case of I think the grammaticalization is evident from its syntactic versatility and its phonologically reduced form (Aijmer 1997). In the case of no doubt there would, if the hypothesis is correct, be a drastic reduction in form as well as increased syntactic versatility.

2.3 I have no doubt

2.3.1 Syntactic patterning
The syntactic complementation of I have no doubt is similar to that of there is no doubt. Its frequency is much lower, as shown in Table 1. Its positional flexibility, however, is greater than that of there is no doubt. It occurs in initial, medial and final position:

(44) I have no doubt that most of our partners now realize that British scepticism about economic and monetary union was right and that Europe must be saved from unthinking, slap-happy federalism. (BNC; AK9, 1681; written/non-fiction/newspaper)

(45) I queued a long time for that loaf and I’ve no doubt it’s the last we’ll be seeing till Thursday. (BNC, AC5, 260; written/fiction)

(46) The children will be delighted with it in the morning, and they’ll have us all up before breakfast, I’ve no doubt. Well, good night, McCann. (BNC, AT7, 1753; written/fiction)

(47) So the Civil Contingencies Unit, the Cabinet committee set up after the miners’ strike of ’72, which I’ve no doubt you were intimately involved in, in ’74 actually delivered the goods. (BNC, B0H, 913; written/non-fiction)
(48) Rascals, I have no doubt, are at work among them. (BNC, EF4, 206; written/non-fiction)

(49) Thereafter she suffered from thrombosed haemorrhoids again I’ve no doubt due to the accident. (ICE-GB, S2A-062/41; spoken/legal presentations)

While the BNC has 772 instances of there is no doubt versus only 20 of there’s no doubt, it has 85 instances of I have no doubt (58%) and 62 of I’ve no doubt (42%). This indicates that ‘I have no doubt’ as a construction occurs in more informal contexts than the construction ‘there is no doubt’. Further, I’ve no doubt is more frequent in non-initial position (53 instances, i.e. 85%) than in initial position (9 instances). In contrast, I have no doubt is initial in 53 instances out of 85, i.e. 62%, versus non-initial in 30 cases, i.e. 38%. This shows that I’ve no doubt is positionally more flexible than I have no doubt. Since the construction ‘I have no doubt’ has greater positional flexibility than ‘there is no doubt’, since it seems to occur more in informal contexts and since the contracted form I’ve no doubt is positionally more flexible than the full form I have no doubt, there appears to be a correlation between informality and positional flexibility. Finally, the subordinator that after initial I’ve no doubt is deleted in all nine cases. Taken together, these factors point to a further development of I have no doubt towards a parenthetical expression than there is no doubt (see Thompson & Mulac 1991 on criteria for the grammaticalization of epistemic parentheticals).

2.3.2 Contextual meanings

In contrast with there is no doubt, which conveys absence of doubt in a generic sense, I have no doubt is explicitly subjective. In some contexts it cannot be substituted by there is no doubt without a significant shift in meaning. One such context is where I have no doubt is a response to the previous speaker’s statement and merely signals polite agreement. Consider example (50):

(50) “Not that we didn’t check on Elsie, as much as we could”.
    “Mrs. Blackler told me that she had worked somewhere else in London”.
    “Only for a few months, yes. She looked after some children in their last year before they were packed off to boarding school. She was given an excellent testimonial”.
    “I’ve no doubt. But did you ask her about her life before she came to London?”
    (BNC, ASN, 778; written/fiction)

Example (50) illustrates how I have no doubt is used to contrast a proposition which the speaker accepts but contrasts with a more important question. This backgrounding function is also typical of no doubt. Further, I have no doubt as a response means ‘I am convinced that you are right, even though I myself have no evidence’. Its meaning is close to I’m sure in this context. There is no doubt would be pragmatically odd in (50) as it would signal that the information is not new to the speaker.

I have no doubt has a subjective orientation and expresses the speaker’s inner conviction, which is not necessarily based on evidence. In cases where it is very clear that
No doubt and related expressions

the speaker cannot have any evidence it is not replaceable by there is no doubt. This is for instance the case in statements about the addressee, which invite confirmation:

(51) “Is there anything you can tell me about the two young women who died after they had been working in this house (...)?”

(...) She said calmly: “Very little, I’m afraid. I’ve no doubt you know most of it already”. (BNC, CJF, 2834; written/fiction)

2.4 A comparison of no doubt, there is no doubt and I have no doubt

Summing up, how do the expressions no doubt, there is no doubt and I have no doubt compare with one another?

In terms of positional flexibility the data show that there is no doubt is the least flexible expression, no doubt the most flexible and I have no doubt is in between the two. In terms of degrees of orientation there is no doubt is basically intersubjective (though it may in some contexts acquire a subjective orientation), I have no doubt is explicitly subjective, and no doubt is neutral in this respect, though it typically expresses a subjective assessment. In terms of degree of epistemic certainty, there is no doubt typically expresses a high degree of certainty, I have no doubt may express certainty or a high degree of probability and no doubt typically expresses probability although it may express certainty in some contexts. I have no doubt and no doubt are both used with a weaker meaning as a response marker indicating ‘I accept what you are saying’, while there is no doubt would be odd in such contexts. The subjectivity of I have no doubt and no doubt makes them into markers which can be thrown into exchanges to agree, to acknowledge information in order to oppose it with some other point, to invite confirmation from the addressee. The analysis seems to indicate that no doubt is pragmatically closer to I have no doubt than to there is no doubt. This does, however, not mean that it developed historically from the former rather than from the latter. Both expressions may have played a role in its emergence.

3. A diachronic excursion

Aijmer (2002a) explains the development of surely into a marker of uncertainty through the process of grammaticalization. Surely developed from an adverb of manner (surely = ‘safely’, ‘securely’) into a sentence adverb and then further into a discourse marker. She points out that the function of surely as a discourse marker is to seek confirmation. It is associated with questioning and uncertainty (Aijmer 2002a:109; see also Downing 2001). A similar explanation seems plausible for no doubt.

The synchronic data have shown that no doubt is used as an epistemic marker (expressing varying degrees of certainty) as well as a discourse marker inviting confir-
Table 4. *No doubt* in the historical corpora

<table>
<thead>
<tr>
<th>Forms</th>
<th>HC</th>
<th>CEECS</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>no doubt</em></td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td><em>no dowte</em></td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td><em>no doute</em></td>
<td>4</td>
<td>–</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>14</td>
<td>4</td>
</tr>
</tbody>
</table>

...mation. In contrast, *there is no doubt* tends to express certainty rather than uncertainty. This would mean that the adverb *no doubt* has acquired the meaning of uncertainty in the course of its development. Since the three expressions continue to exist side by side they have developed specialized meanings. *There is no doubt* is explicitly objective (in Halliday's sense it is an objectifying grammatical metaphor of modality, Halliday 1994:354–355), while *I have no doubt* is explicitly subjective. *No doubt* has no explicit indicator of objectivity or subjectivity and is used in both types of contexts. When it functions as a discourse marker, expressing agreement or inviting confirmation, it is subjective and conveys probability or predictability but not certainty.

For diachronic corpus evidence of earlier uses and meanings of the expressions I consulted the Helsinki Corpus (HC) and the Corpus of Early English Correspondence Sampler (CEECS).

Table 4 gives the frequencies of *no doubt* in various spellings in HC and CEECS.

The earliest occurrence is the form *no doute*, in Purvey’s *Prologue to the bible*, a religious treatise from the period 1350–1420. All other occurrences date from early Modern English (1500–1570) and later. In the example from Purvey *no doute* has initial position:

(52) and I preie, for charite and for comoun profyt of cristene soulis, that if ony wiys man fynde ony defaute of the truthe of translacioun, let him sette in the trewe sentence and opin of holi writ, but loke that he examyne truli his Latyn bible, for *no doute* he shal fynde ful manye biblis in Latyn ful false, if he loke manie, nameli newe; and the comune Latyn biblis han more nede to be correctid, as manie as I haue seen in my lif, than hath the English bible late translatid; and where the Ebru, bi witnesse of Jerom, of Lire, and othere expositouris discordith fro oure Latyn biblis, I haue set in the margyn, bi manner of a glose, what the Ebru hath (HC, Purvey, The Prologue to the Bible. Religious treatises (1350–1420))

The later occurrences of *no doubt* are in initial, medial and final position. Here are some examples:

(53) invincible, like stones, which abide in the building for ever and fall not out. No, it is not the worthinesse of our beleuving, it is the vertue of him in whom we beleue, by which we stand sure as houses that are buildeid vpon a rocke. He is a wise man, which hath buildeid his house vpon a rocke; for he hath chosen a good foundation, and *no doubt* his house will stand. But how shall
it stand? verily by the strength of the rocke which beareth it, and by nothing else. Our fathers, whom God delivered out of the land of Egypt, were a people, that had no peeres amongst the nations of the earth, because they were built by faith vpon the rocke, which rocke is Christ. (HC, Hooker, Two Sermons upon Part of S. Judes Epistle. Epistles (1570–1640))

(54) so that in reason, as wel as in experence, there fal out to be these 3. distempers (as I may tearm them) of learning: The first fantastical learning: The second contentious learning, and the last delicate learning, vaine Imaginations, vaine Altercations, and vain affectations: and with the last I wil begin, (^Martin Luther^) conducted (no doubt) by an higher prouidence, but in discourse of reason, finding what a Prouince he had vndertaken against the Bishop of (^Rome^), and the degenerate traditions of the Church, and finding his owne solitude, being no waies ayded by the opinions of his owne time, was enforced to awake all Antiquitie, and to call former times to his succors, to make a partie against the present time: so that the ancient Authors (HC, Bacon, The Twoo Bookes. . . Advancement of Learning. Educational treatises (1570–1640))

(55) Ah! sayes hee, this must haue, must haue a good showre to clense it; and with that goes behinde the arras. Harr y, saies hee, Ile foe behind the arras, and study three questions, and come againe; see, therefore, you lay aside this melancloly muse, and study to answere me. I, quoth the king: they will be wise ones, no doubt. At last out comes William with his wit, as the foole of the play does, with an anticke looke to please the # beholders. Harry, sayes hee, what is it, that the lesser it is, the more # it is to be feared? The king mused at it; but, to grace the jest better, he answered, he knew not. Will answered, it was a little bridge ouer a deepe riuer; at which hee smyled. (Armin, A Nest of Ninnies. Fiction (1570–1640))

In Early Modern English the expression seems to have become well established. It is sometimes inserted parenthetically. The extent to which the meaning is ‘certain’ or ‘very probable’ is very difficult to assess and it is not always possible to do so. In some cases, however, one interpretation is more plausible than another one. There are instances in which no doubt expresses certainty rather than uncertainty, as in (52) above. However, all instances in HC and CEECS occur in contexts in which there is always room for error or disagreement: pronouncements about the future (with will, shall, may, might), but also about the past and present. Hence, the shift towards probability is not a radical one.

It is interesting to note that the form saunz doute occurs still earlier, in the romance Kyng Alisaunder (period 1250–1350). The HC gives three occurrences of the French expression in this text. In all three cases saunz doute provides a rhyme for the words route and aboute. The meaning of the Old French expression was ‘without fear’ or ‘without any doubt’ (see further Section 4). Here is an example:
(56) Þe erþe quaked of her rydyng; Þe weder þicked of her crieyng. Þe blood of hem þat weren yslawe Ran by flodes to þe lowe, And Jou sigge, sikerlich, Darrie fuerst wel douršttilich, And dude swiþe mychel woo. To on syde he drouþe hym þoo – He blew an horne quyk, saunz doute. His folk hym com swiþe aboute, And hem he seide, wiþ voice clere: 'Jch bidde, frendes, þat þe me here! Alisaunder is comen in þis londe, Wijþ stronge kniþtes, wiþ miþty honde. 3if he passeþ wiþ honoure, Oure is al þe dishonoure. (HC, Kyng Alisaunder. Romance (1250–1350))

It may be no coincidence that the present-day French adverb sans doute, like English no doubt, expresses probability rather than certainty (cf. e.g. Lindvall 2004). The French expression may, in other words, have influenced the development of the probability meaning of no doubt, though this is difficult to prove. Alternatively, the French and the English expressions may have had a parallel semantic development.

The existential expression there is no doubt occurs earlier in the HC than the short form no doubt. The following two instances are both from the period 1250–1350:

(57) King Grander was of herte grim and rod to Beues and he to him; And ase þei boþe to gedre mete, Wiþ here launces þei gonne mete, Þat hit gonnen al to-driue and te-borsten on pises fiue. Here swerdes drowe kniþtes stoute And fiþte faste, it is no doute; Þe medwe squarþte of her dentes, Þe fur fleþ out, so spark o flintes; Þus þai leide on in boþe side Be-twene midmorwe and vndertide. King Grander was agremed strong, Þat sire Beues him stod so long, And wiþ is swerd a hitte is scheld, A quarter fel in to þe feld, Hauber, plate and aktoun, In to Beues forþer arsoun Half a fot he karf doun riþt. Po Beues seþ þat strok of miþt (HC, The Romance of Sir Beues of Hamtoun. Romances (1250–1350))

(58) Sir simond de mountfort . þo ido was al þis. Vorþ mid him þe king huld . as in warde ivis . and þe king of alemaine . and sir edward al so. In þe castel of walingford . in warde he leet do . []. HENRICUS .] and oþer men þat were inome . he let bringe aboute . In oþer castels vaste inou . þat þe nes no doute. (Robert of Gloucester, History (1250–1350))

Later instances are for example found in Chaucer’s Boethius, which is argumentative discourse:

(59) “Ryght so is it,” quod I. “Suffisaunce and power ben thanne of o kynde?” “So semeth it,” quod I. “And demestow,” quod sche, “that a thyng that is of this manere ((^that is to seyn, suffisaunt and mighty^)) oughte ben despised, or ellis that it be right digne of reverence aboven alle thynges?” “Certes,” quod I, “it nys no doute that it nys right worthy to ben reverenced”. “Lat us,” quod sche, “adden thanne reverence to suffisaunce and to power, so that we demen that thise thre thynges be al o thyng. “Certes,” quod I, “lat us adden it, yif we wil graunten the sothe”. (Chaucer, Boethius. Philosophy (1350–1420))
No doubt and related expressions

It will be noted that the expression had positional flexibility, which would have facilitated the development of the versatile adverb *no doubt*.

The earliest occurrence of the expression *I have no doubt* in the HC is in the period 1420–1500 and has clause-final position:

(60) Now trewly my lorde þe kynge we had ben vn-hende and nevyr non of us Able for to be a knyght If þat Any of us to hem had ben a frende and asavyd Any lyf þæn þi mekyl myght Ffrom deth hem to flytt. (Æherodes (†rex)*)

Amonges all þat grett rowthte he is ded *I have no doubt* perfore menstrell rownd a-bowte blowe up a mery fytt. (Ludus Coventriae. Drama, mystery plays (1420–1500))

It is also used in imperative clauses with a second person subject, as in (61):

(61) But, sir,’ # seyde the lady, ‘as thou arte called the worshypfullyest knyght of the # worlde, I requyre the of trewe knyghthode, kepe me and save me, for # whatsomever he sey he woll sle me, for he is withoute mercy.’ *Have ye no doute*: hit shalle nat lye in his power. ‘Sir,’ seyde the knyght, ‘in your syght I woll be ruled as # ye woll have me.’ And so sir Launcelot rode on the one syde and she on the # other syde. And he had nat redyn but a whyle but the knyght bade sir Launcelot turne hym and loke behynde hym, and seyde, ‘Sir, # yondir com men of armys aftir [[us]] rydyenge.’ (Malory, Morte Dartur. Romance (1420–1500))

The imperative form, used to encourage the addressee not to have any doubts, is at the same time a signal of the speaker’s certainty. What such instances show is that the interactional uses of *have no doubt* exploit the rhetorical potential of the expression.

It will be noted that the earliest occurrence of *have no doubt* is later in HC than the existential structure. However, this is a coincidence, as shown by the examples in the electronic *Middle English Dictionary*:

(62) Denemark shal be þin..Haue þou nouth þer-offe douthe! (c1300. Havelok)

(63) For al the good..He lese shal; ther of haue I no doute. (c1395. Chaucer C.T.)

A final remark is that the word *doubt* occurs much more frequently in the two historical corpora in various negative expressions than in positive ones (see Table 5).

Speakers have, it appears, more need to say that they have no doubt about something than the opposite. They use such expressions to convey either that they as speakers ‘have no doubt’ or that the addressees should ‘have no doubt’. From a dia-

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Positive</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>83 (44%)</td>
<td>106 (56%)</td>
<td>189</td>
</tr>
<tr>
<td>CEECS</td>
<td>81 (40%)</td>
<td>123 (60%)</td>
<td>204</td>
</tr>
</tbody>
</table>
logic point of view negation presupposes an alternative positive (cf. White 2003). As such these negative expressions are examples ‘par excellence’ of the rhetorical function of certainty expressions, to position the speaker vis-à-vis other viewpoints in a heteroglossic world.

4. Discussion

How do the synchronic and diachronic findings provide an answer to the question of how we can explain that no doubt does not express certainty and that in general the meanings of certainty and uncertainty are closely related?

There are two factors which are important here. One is the contexts in which epistemic expressions of certainty are used, the other is the emergence of a new adverb from clausal expressions. These two factors play a role in the development of a weaker degree of certainty in no doubt.

The contexts in which certainty markers occur are those in which speakers feel the need to emphasize their commitment to the truth value of their statements. Since speakers are normally assumed to believe that all their statements are true, adding certainty expressions is marked and leads to inferencing (according to the Gricean maxim ‘Say no more than you must’ (1975)). Speakers may have various reasons to emphasize their commitment, which all have to do with their dialogic positioning within a heteroglossic world. Following White (2003), we should explain the use of modality not only from a truth-functional point of view (‘How certain or uncertain is the speaker about the truth-value of his or her statements?’) but from a rhetorical point of view (‘How does the speaker position his or her utterance vis-à-vis other real or imagined utterances?’). Certainty expressions are used in contexts where speakers cannot take a common world view for granted. This means that the statement thus qualified as ‘certain’ is either disputed or conceded to be undisputed in contrast with some other, disputed statement. This explains why certainty markers often occur in concessive contexts. It also accounts for the use of certainty expressions in cases where speakers express inner beliefs in the truth value without having evidence. All such ‘heteroglossic’ contexts create pragmatic meanings of adverbs of certainty which through ‘pragmatic strengthening’ (Traugott 1989) may become conventionalized. In the case of no doubt the inference that there is in fact some reason for doubt has become conventionalized. However, the corpus data showed that in some cases, notably in concessive contexts, no doubt seems to convey certainty rather than probability. Example (16) was a case in point. A parallel can be found here with the use of may in concessive clauses (He may be good-looking but he’s not a likeable person), which means ‘I grant you the truth of P’. May can be used in such contexts even when the speaker has no doubts about P. However, this may can still be seen as a rhetorical ploy, expressing some sort of ‘reluctant concession’ in order to weaken commitment and present the statement as ‘not of crucial importance’. The same applies to no doubt.
An account of the close link between certainty and uncertainty in heteroglossic terms does not, however, explain why *no doubt* expresses a weaker degree of epistemic certainty than *there is no doubt* or even *I have no doubt* in some contexts. The explanation lies in the reduction of the form and the development of discourse marker characteristics by the adverb *no doubt*. First, in contrast with the clausal expressions, *no doubt* has no surface indicators of intersubjectivity or subjectivity. This makes it more flexible in this respect and hence rhetorically useful. Second, its reduced form makes it less prominent and hence a more likely candidate for the expression of tentative rather than assertive statements. In this connection it is interesting to note that the ‘fuller’ form *no doubt about it* expresses a higher degree of certainty than *no doubt* (see e.g. Longman 1995; Macmillan 2002). Third, because of its reduced form it can easily be inserted in all positions in the clause, and hence qualify very subtly various aspects of a statement. This flexibility and lack of prominence make it into a useful marker of expressive meanings (such as irony) in the conversational give-and-take.

The diachronic data show that expressions referring to absence of doubt are more frequent than those expressing doubt. This in itself is indicative of the rhetorical reasons speakers have for conveying epistemic certainty. It further appears that those negative expressions not only convey the speaker’s certainty but also tell the addressee to be certain. The interactional use of these expressions may have led to the grammaticalization process of *no doubt*. Since both *there is no doubt* and *I have no doubt* are also in use today, they have their own specialized uses, different from *no doubt*. That *there is no doubt* is much more frequent than *I have no doubt* may be explicable from its rhetorical usefulness as an ‘objectifying metaphor’ of modality, in the sense that it represents the speaker’s subjective assessment as if it were intersubjective (Halliday 1994:362). It was also shown that *I have no doubt* is closer to *no doubt* in various uses, which could explain the lesser degree of the former in interactional usefulness.

Finally, the semantic development of the word *doubt* shows that the expression followed the predictable path of all epistemic expressions (Traugott 1989), i.e. from ideational to interpersonal and textual meanings. Middle English borrowed the word *doute* from French, where it meant ‘fear’ as well as ‘hesitation’ and ‘doubt’ (Philippe Verelst, p.c.). According to the electronic *Middle English Dictionary* earlier meanings of *doubt* were indeed also ‘anxiety, fear, fright’ and ‘a cause or reason for fear, something to be feared, danger, peril’. Thus *haven doute* meant ‘be afraid’. The meanings ‘to hesitate’ and ‘to be uncertain’ are linked to ‘be afraid’ in very obvious ways. While ‘fear’ and ‘hesitation’ apply to future events and actions, the epistemic sense of ‘uncertainty’ may also apply to present and past events. Such a development has also taken place in the case of *I’m afraid*, which has completely grammaticalized in instances such as *I’m afraid he died last year* (see Kjellmer 2003). In both cases the meaning of ‘fear’ has been lost. While *I’m afraid* has developed into a politeness marker, *no doubt* has acquired epistemic meaning as well as discourse functions such as agreement.

To conclude I would like to return to two questions posed at the outset which I see as crucial to the issues discussed here. These are the questions on the status of the meaning of uncertainty in *no doubt* and on the role of rhetorical factors in an account...
of the relationship between certainty and uncertainty as conveyed by one and the same
linguistic expression. It seems to me that these two questions are related. The answer to
the first question, viz. whether the meaning of ‘lack of complete certainty’ is part of
the coded meaning of no doubt, is positive, and the explanation is the rhetorical exploita-
tion of the expressions. What seems to have taken place is a development from the
conceptual meaning of ‘absence of doubt’ to the epistemic meaning of certainty on the
part of the speaker. Epistemic meaning is necessarily subjective, i.e. representing the
speaker’s viewpoint. In that sense we have a development towards greater subjectivity.
Now, as pointed out by White (2003), speakers use epistemic expressions not only and
perhaps not even primarily to convey the state of their knowledge but also to position
themselves vis-à-vis other speakers. Expressions of certainty are especially useful
in contexts of argumentation, where there is divergence of opinion, to emphasize one’s
own viewpoint. In such contexts the next step will be the use of certainty expressions
even when the speaker does not have absolute certainty but wishes to convey it to per-
suade or to reassure, i.e. for interpersonal reasons. Traugott and Dasher (2002:24),
referring to Halliday (1994:362–363), point out that for argumentative reasons “subjective
opinions which are in fact best expressed by e.g. I think, are concealed behind
objective-seeming expressions like surely”. We can add that as a result adverbs such as
surely will tend to lose their meaning of certainty. If such meanings arise frequently
in particular contexts the inference that the speaker is in fact not certain may become
conventionalized and lead to a new coded meaning. This is the development sketched
by Traugott and Dasher (2002:35), viz. the path from coded meanings to utterance-
token meanings to utterance-type meanings to new semantically polysemous (coded)
meanings. The further development of no doubt as a discourse marker with interper-
sonal, procedural meaning has been shown to have taken place. No doubt can be used
to invite confirmation, or to express sarcasm and irony. These are purely intersubjec-
tive meanings (in the sense here of Traugott and Dasher (2002:23), as distinct from
Nuyts’ use of the term, as explained in Section 2.2.2 above). Again, the path from sub-
jectivity to intersubjectivity is a development that must be explained from rhetorical
uses to which expressions are put (Traugott & Dasher 2002:24). This article has thus
attempted to contribute further evidence for the need to study modal expressions in
dialogic terms, as means for negotiating opinions.

The study has also shown that there are indications that the development of the
adverb no doubt is the result of grammaticalization. Semantically, there is a clear
development from an ideational meaning of the word doubt (‘fear’, ‘anxiety’) to an
epistemic one. Grammatically, a nominal group developed into an adverb with great
positional flexibility. It is plausible to assume that the nominal group is a shortened
version of the clausal construction ‘there is no doubt’, which not only occurred earlier
but which has also retained its sense of ‘certainty’. The adverb seems to have devel-
oped the weaker sense of probability, most likely as a result of its frequent use for
rhetorical purposes in contexts of ‘uncertainty’. The further development of the over-
tone of irony when it functions as a discourse marker points to the accumulation of
pragmatic meanings which accompanies the process of intersubjectification, follow-
No doubt and related expressions

ing subjectification. Other developments of this kind have been described by Traugott and Dasher (2002:152ff.), while Byloo et al. (this volume) produce evidence for the rise of the Dutch discourse particle zeker out of the weaker epistemic use of the adverb. The case they describe in this article has many parallels with that of no doubt. The historical data seem to suggest, however, that several factors have played a role in the development of the adverb no doubt. Both the syntactic versatility of there is no doubt and I have no doubt in the historical corpora and the semantic-pragmatic relatedness of certainty and probability are important. Finally, the parallel with the semantic-pragmatic development of the French expression sans doute is striking and deserves further examination. A more thorough historical study is called for to test the hypotheses.

References

Middle English Dictionary. Electronic version: http://www.hti.umich.edu/c/cme/


On *certainly* and *zeker*

Pieter Byloo, Richard Kastein and Jan Nuyts
University of Antwerp

This contribution investigates the meanings and uses of the English and Dutch adverbs *certainly* and *zeker*. In the range of adverbial expressions of epistemic modality, the 'strong' adverbs (expressing 'certainty') clearly tend to be used in a much more 'flexible' way than the (relatively) weaker ones, and this contribution is an attempt to get a better view of what this flexibility actually involves, on the basis of a careful analysis of corpus data. The comparative approach allows us to distinguish between features which are whims of a single language, and hence have little meaning beyond the grammar of that language, and features which reflect more profound issues in the analysis of the meanings at stake, and which therefore may reflect fundamental conceptual issues.

1. Introduction

This paper investigates the meanings and uses of the (British) English and Dutch adverbs *certainly* and *zeker*. This case study is part of a wider project (see Note 1) aiming to achieve a better understanding of the conceptual status of modal categories, in elaboration of the preparatory investigations in this regard in Nuyts (2001). In the range of adverbial expressions of epistemic modality, the analysis in Nuyts (2001) focused on the adverb *probably* and its Dutch and German equivalents. Cursory observations indicate, however, that the stronger epistemic adverbs tend to be used in a much more 'flexible' way, and the present paper is an attempt to get a better view of what this flexibility actually involves, on the basis of a careful analysis of corpus data. The comparative approach, as a minimal precursor to a wider, truly typological analysis, may help to distinguish between features of the forms under scrutiny which are whims of a single language, hence have little meaning beyond the grammar of that language, and

---

1. This research is part of a ‘VNC project’ on ‘Modal auxiliaries and other expressions of modality in Dutch’ (FWO-Flanders and NWO) and a ‘GOA’ on ‘Mood and modality’ (Research Council of the University of Antwerp). The 6th release of the ‘Corpus Gesproken Nederlands’ was made available by the ‘Nederlandse taalunie’ and the European Language Resources Distribution Agency (ELDA) to Jan Nuyts, as a member of the ‘CGN user group’.
features which reflect more profound issues in the analysis of the meanings at stake, and which therefore may reflect fundamental conceptual issues.

This paper is organized as follows: Section 2 introduces the issue and briefly reviews some known facts and some earlier literature on these adverbs. In Section 3 we present our data and Section 4 offers an analysis. Conclusions are presented in Section 5.

2. What we do and do not know about certainly and zeker

Let’s start where everything starts: in the past. A combination of information from different sources (Kurath 1959; Murray et al. 1933–1970; Onions 1966; Partridge 1966 for English; Pijnenburg et al. 2001; Verwijs & Verdam 1885–1929; De Vries & Te Winkel 1864–1998; Van Wijk 1912 for Dutch) indicates that certainly and zeker have a comparable though not identical diachrony. Both have Latin origins, but different ones. The English adverb, or the corresponding adjective certain, is a loan from Late Old French certain, which is in turn a development, via Vulgar Latin *certanus, from Latin certus “decided, determined, sure”. The Dutch form is a very old loan (witness the sound changes in the German cognate sicher) from Latin securus, probably originally adopted as a legal term used to discharge a person from any verdicts. These facts are compatible with the hypothesis, emerging from current views about paths of meaning change from more ‘objective’ to more ‘subjective’ (e.g. Traugott & Dasher 2002), that these forms started out as ‘manner adjectives/adverbs’, a type of meaning they still express today, even if rather marginally (see Section 3). Our sources offer no explicit proof for this, however, since they indicate that the epistemic meaning – manifestly present today – was present from the earliest Middle English and Middle Dutch attestations onwards (according to the Oxford English Dictionary (Murray et al. 1933–1970) even Latin certus already had all the meanings found in Middle English). Also, our sources do not allow us to determine whether the adjectival or the adverbial usage is the older one: both were fully present in Middle English and Middle Dutch, with both the manner and epistemic meanings (although, remarkably, the latter appears still fairly marginal for the adjectival use, but not for the adverbial use, in Middle Dutch).

For adjectival certain and zeker, the foregoing by and large exhausts their meaning evolution. But for adverbial certainly and zeker, this is not the end of the story: in present day English and Dutch, these forms are used in ways which cannot be categorized as ‘marking manner’ or ‘expressing an epistemic judgment’. (There are actually traces of some of these other uses in our diachronic sources, even if most of them are not clearly identified there – we return to these later.) These other uses have received fairly little explicit attention in the literature, however.

Thus, English certainly is usually characterized as an epistemic adverb (e.g. Merlini Barbaresi 1987; Hoye 1997; Ramat & Ricca 1998; Nuyts 2001), and otherwise little is to be found beyond cursory observations that it is also used, for instance as an intensifier (Greenbaum 1969) or an emphasisizer (Bolinger 1972; Hoye 1997).
Dutch zeker has received a little more attention (see Van der Wouden 2000 for a literature survey). This form, too, is usually treated as an epistemic adverb (cf. e.g. Nuyts 2001), albeit a fairly complex one (cf. Dirven 1973; Elffers 2000; see Section 4.1). But in a more systematic corpus-based analysis, Van der Wouden (2000) also describes uses of zeker beyond the epistemic one. He treats zeker in these other uses as a particle, and – in the light of the literature on particles – he identifies uses as a modal particle, a focus particle and a scalar particle. Van der Wouden adopts the traditional definition of this distinction in syntactic terms, however, and consequently his overall classification of uses of zeker consists of a mixture of semantic and (often quite superficial) formal categories. Moreover, he usually characterizes the semantic dimensions in terms of paraphrases rather than definitions, which is not unproblematic (see Section 4).

In this paper, then, we will attempt to throw more light on the range of what we will call ‘post-manner adverbial uses’ of certainly and zeker, i.e. all those uses of the adverbs which have evolved beyond the (probably) oldest manner use.

3. The data: Sources and selection criteria

We will base our investigation on corpus data from two types of sources:

i. In order to be able to compare the English and Dutch form directly, we looked at their occurrences in Harry Potter and the Chamber of Secrets (Rowling 1998) and its Dutch translation (Rowling 1999). This yielded 9 relevant (see below) instances of certainly and 22 relevant instances of zeker (including jazeker). In only 4 cases do certainly and zeker (including one case of jazeker) correspond in the English and Dutch versions.

ii. To obtain more data, we randomly selected 100 instances of certainly from the spontaneous spoken parts of the British National Corpus (BNC) World Version (December 2000), and 100 instances of adverbial zeker – including jazeker – from the 6th release (November 2002) of the ‘Corpus Gesproken Nederlands’ (CGN, a corpus of spoken Dutch), excluding the parts with ‘read aloud’ data. Zeker and jazeker had to be selected separately, but we have respected their estimated relative frequency in the CGN overall; consequently, our sample includes 98 instances of zeker and 2 of jazeker.

Since we are interested in ‘post-manner adverbial uses’ of certainly and zeker, we excluded from our sample (from both sources) the following:

– the negative forms uncertainly (only used as a manner adverb) and onzeker (only used as a manner adverb or a manner or epistemic adjective);

2. We have included this form in the investigation since it is not really different from the combination ja(,) zeker, which is within our range of relevant data (see below).
– uses of certainly and zeker as a manner adverb;
– adjectival uses of zeker;
– uses of zeker as an indefinite determiner, as in *een zekere Hendrik* “a certain Hendrik”;
– uses of zeker in *zeker weten* “know for sure” and *langzaam maar zeker* “slowly but surely”: these idioms are so strongly ‘fossilized’ that it is nearly impossible to determine the role of zeker in them (it might even be a manner use).

We did include ‘absolutive’ uses (Nuyts 2001:89) of Dutch zeker (*jazeker* is even always used in an absolutive way), in which it functions ‘autonomously’ as a reaction to a question or remark by the interlocutor, of the kind in (1): in Dutch one cannot formally decide whether these are adverbial or adjectival uses (in the latter case we would be dealing with an elliptical variant of the pattern *dat is zeker* “that is certain”), but since English systematically uses an adverb in such circumstances, we assume the same for Dutch.

(1) A: *En ben je nu blij* “Are you happy now?”
B: *(Ja,)* zeker. “(Yes,) certainly”.

We have obviously excluded uninterpretable instances of certainly and zeker (e.g. when they appear in incomplete or interrupted utterances, when there is insufficient verbal or non-verbal context to make them comprehensible, etc.).

Since many of the instances are ambiguous, in the analyses below we will distinguish between minimal and maximal counts: the minimal counts only include unambiguous cases of a category, while the maximal counts also include the ambiguous instances.

4. The analysis

The Harry Potter data invite at least two preliminary remarks:

i. The fact that Dutch zeker occurs nearly two and a half times as often as English certainly suggests that the former is used much more systematically – hence possibly in more different ways – than the latter. Maybe the existence of surely as a 'com-
peting’ form in English but the absence of such a competitor in Dutch explains this difference.

ii. The fact that certainly and zeker appear in corresponding places in only 4 instances in the English and Dutch texts might be taken to suggest that the two forms are far from equivalent in their meanings and uses.

A closer scrutiny of both the Potter and the BNC/CGN cases reveals, however, that (ii) is only partially true.5 There appear to be 4 major types of uses in the data, 3 of which are shared by the English and Dutch forms – only one type is manifestly present for the Dutch form but at best marginally, and possibly not at all, for the English form. (As will appear in the detailed analyses below, however, within the shared uses some further differences between the two languages can be observed.)

Here is a brief preliminary definition of the four uses – the ‘pragmatic’ use is possibly missing in English:

– the ‘epistemic’ use: the adverb expresses the degree of likelihood of a state of affairs;
– the ‘scalar’ use: the adverb situates (an entity in) a state of affairs on a scale or in a range of conceivable elements (related entities or states of affairs);
– the ‘strengthening’ use: the adverb reinforces an assessment of a state of affairs;
– the ‘pragmatic’ use: the adverb serves as a speech act modifier turning a declarative into a special kind of interrogative.

Detailed definitions and illustrations of these categories will be provided below. Table 1 shows the relative share, in minimal and maximal count (see Section 3),6 of these uses in the total number of instances of the two forms in the Potter data, the BNC/CGN data, and in total. Table 2 lists the types of ambiguity and their absolute frequency.7 (Actually, the 4 corresponding instances in the English and Dutch versions of Potter involve 3 strengthening cases and one epistemic case.)

Before going through the different meanings or uses in more detail, a few words about our method for classifying instances. The classification is exclusively based on semantic criteria, viz. on determining whether the meaning of an instance is in accordance with the definition of a category.8 We have not used other analytical ‘tools’ –

5. In fact, most instances of certainly and zeker that are now rendered by other (often semantically completely unrelated) expressions in the other language version of Potter could perfectly well have been rendered by means of the corresponding adverb.

6. As a consequence of this way of counting, neither the minimal nor the maximal (relative) totals for each corpus category add up to 100.

7. Instances were coded as ambiguous whenever two readings were feasible, even if one of them was more likely.

8. This has obviously been an ‘accumulative’ process. We have started ‘inductively’, without any preconceptions about categories defined elsewhere in the literature, by looking at instances and finding similarities and differences between them. From there we have gradually come to
Pieter Byloo, Richard Kastein and Jan Nuyts

| Table 1. Share of the different meanings/uses of *certainly* and *zeker* in Potter and BNC |
|-----------------------------------------------|--------|--------|-----------------------------------------------|--------|--------|
|                                              | Potter (n = 9) | BNC (n = 100) | sum (n = 109) | Potter (n = 22) | CGN (n = 100) | sum (n = 122) |
|                                              | min (max)      | min (max)     | min (max)     | min (max)       | min (max)     | min (max)     |
| epistemic                                    | 11.1 (11.1)    | 08.0 (18.0)   | 08.3 (17.4)   | 13.6 (13.6)     | 12.0 (21.0)   | 12.3 (19.7)   |
| scalar                                       | 00.0 (22.2)    | 24.0 (45.0)   | 22.0 (43.1)   | 18.2 (18.2)     | 33.0 (35.0)   | 30.3 (32.0)   |
| strengthening                                 | 66.7 (88.9)    | 39.0 (65.0)   | 41.3 (67.0)   | 27.3 (27.3)     | 25.0 (32.0)   | 25.4 (31.1)   |
| pragmatic                                    | 00.0 (00.0)    | 00.0 (02.0)   | 00.0 (01.8)   | 40.9 (40.9)     | 20.0 (22.0)   | 23.8 (25.4)   |

| Table 2. Absolute frequency of ambiguous cases of *certainly* and *zeker* in Potter and BNC |
|-----------------------------------------------|--------|--------|-----------------------------------------------|--------|--------|
|                                              | Potter (n = 9) | BNC (n = 100) | sum (n = 109) | Potter (n = 22) | CGN (n = 100) | sum (n = 122) |
|                                              | min (max)      | min (max)     | min (max)     | min (max)       | min (max)     | min (max)     |
| epistemic/scalar                             | 0        | 3        | 3                | 0        | 2        | 2                |
| epistemic/strengthening                      | 0        | 6        | 6                | 0        | 6        | 6                |
| epistemic/pragmatic                          | 0        | 0        | 0                | 0        | 1        | 1                |
| scalar/strengthening                         | 2        | 17       | 19               | 0        | 0        | 0                |
| scalar/pragmatic                             | 0        | 0        | 0                | 0        | 0        | 0                |
| strengthening/pragmatic                      | 0        | 2        | 2                | 0        | 1        | 1                |
| epist./scalar/strengthen.                    | 0        | 1        | 1                | 0        | 0        | 0                |

specifically, unlike Van der Wouden (2000), we have not used paraphrases to help catego- 

erize cases. The major reason is that the meaning of the forms which might be used in 

the paraphrases is usually as undefined (in terms of a detailed and subtle empirical 

semantics of the kind needed here) as is the meaning of *zeker* and *certainly*. And it is 

usually at least as complex and multifacetted.9 The use of such forms ‘at face value’ 

thus leads straight into a methodological quagmire.

4.1 The epistemic use

The epistemic use of *certainly* and *zeker* obviously involves an expression of the speaker’s estimation of the likelihood of the state of affairs. As already suggested in 

the 4 categories as defined in this paper. So we have constantly gone back and forth between data 

and definitions.

9. For instance, van der Wouden (2000) paraphrases one of his ‘modal particle’ uses of *zeker* 

by *in elk geval* “in any case” and *alleszins* “in every respect” (literally, but this does not catch 

the actual meaning of the Dutch form, which is very hard to render in English). However, attempts 

to use these paraphrases in our analyses soon reveal that (i) both of them match some epistemic 

as well as some scalar uses, (ii) neither of them matches all of these uses, and (iii) in cases in 

which there is a match, it rarely feels like a complete one.
Section 2, however, in Dutch – but not in English – this category is more complex than might appear at first sight.

In all English cases, and in the majority of Dutch cases, the adverb expresses certainty that the state of affairs does – or, when combined with negation, does not – apply. Of course, the fact that the speaker uses this adverb instead of an unqualified assertion may sometimes indicate that conceptually (s)he is not really completely certain, but that is rather a pragmatic by-effect of this use. Examples (2) and (3) are from Potter (they involve an instance in which the adverbs correspond in the English and Dutch versions). Examples (4) and (5) are from the BNC and the CGN.10

(2) Harry […] looked up at Justin, grinning, expecting to see Justin looking relieved, or puzzled, or even grateful – but certainly not angry and scared.

(3) Hij keek grijnzend naar Joost, in de verwachting dat die opgelucht of verbaasd of zelfs dankbaar zou zijn – maar zeker niet woedend en bang. ‘Waar denk je dat je mee bezig bent?’ schreeuwde hij [Joost] […] . (PD: 145)

(4) A: Does that hurt there? B: No. A: Not sure what’s causing that […] Well yes I mean it’s more likely to have been a muscle than anything else, with pain that comes and goes. Certainly won’t be a bone that’s causing it. (BNC: G5T 80)

(5) Door het feit dat we hier ja geantwoord hebben kunnen we dus zeker niet hier geraken. (CGN: fv400043)

“Due to the fact that we have answered yes here, we certainly cannot get there”.

As Dirven (1973) and Elffers (2000) have already pointed out, Dutch zeker can also express a weaker epistemic value, albeit one which comes closer to (high) probability. This meaning appears to be absent in English certainly.11 It is not very frequent in Dutch either. However, we found 2 clear cases in the Potter data, (6) and (7), and 2 instances in the CGN, which are both ambiguous between this weaker and the stronger epistemic reading, including (8) (if we include the latter two, this represents 26.7% of the minimal and 16.7% of the maximal number of epistemic cases in the Dutch data). (7) and (8) feature the collocation vast en zeker (literally) “surely and certainly”.12

---

10. All examples are identified by a code: ‘PE’ (i.e. Potter English), ‘PD’ (Potter Dutch), BNC or CGN – PE and PD are accompanied by the page number in the printed books, BNC and CGN by the name of the transcript in these corpora from which the example is drawn, BNC moreover by the position within that transcript. Translations of PD instances are from PE, translations of all other Dutch instances are our own. For the sake of presentation, we have ‘cleaned up’ corpus instances by omitting (without acknowledgment) repetitions, false starts, pause markings, backchannel cues, etc. to the extent that these are irrelevant for the interpretation.

11. But not in English no doubt (cf. Simon-Vandenbergen this volume).

12. This collocation can also have other uses, though – see below.
(6) [Dobby begon] zichzelf met Harry's bureaulamp op het hoofd te slaan. Beneden viel een stilte. Twee tellen later hoorde Harry [...] oom Herman de gang opstormen. 'Dirk heeft zeker de televisie laten aanstaan, de kleine rakker!' riep hij. (PD: 17)

"[Dobby] seized Harry’s desk lamp and started beating himself around the head [...] . A sudden silence fell downstairs. Two seconds later Harry [...] heard Uncle Vernon coming into the hall, calling, ‘Dudley must have left his television on again, the little tyke!’”

(7) Harry en Ron wachtten met ingehouden adem: Malfidus stond vast en zeker op het punt om te zeggen dat hij het was. Maar toen – ‘Ik wou dat ik wist wie het was,’ zei Malfidus kribbig. (PD: 167)

“Harry and Ron waited with bated breath: Malfoy was surely seconds away from telling them it was him. But then – ‘I wish I knew who it is,’ said Malfoy petulantly”.

(8) De eerste oplage van vijftienduizend stuks is in een uur verkocht waarmee het nummer vast en zeker binnenkomt op nummer één. (CGN: fn003037)

“The first edition of fifteen thousand was sold within an hour, hence the song will (most) probably/certainly enter at number one”.

Diachronic sources show traces of this weaker usage (along with the stronger usage) from Middle Dutch onwards, witness (9) (from Verwijs & Verdam 1885–1929) and (10) (taken from Van der Wouden 2000:74 – similar cases can be found in De Vries & Te Winkel 1864–1998).

(9) Si es op ene ander vaert: si mint zeker enen jonghen man. (Esmoreit, app. 1410)

“She is heading somewhere else: she is probably/surely in love with a young man”.

(10) Een windvlaag ontrukte onverwachts den vierkanten lap [...] aan den man, die hem ophield, en voerde dien, eer deze het verhoeden kon, midden in de stroom. De verliezer, zeker veel waarde aan dit voorwerp hechtende, aarzelde niet het na te springen [...]. (J. van Lennep, Onze voorouders in verschillende taferelen geschetst, 1838–1844)

“A gust of wind unexpectedly snatched the square cloth from the man, who held it high, and carried it, before he could prevent this, into the middle of the river. The loser, probably/surely caring a lot about it, did not hesitate to leap after it”.

Instances such as (6) and (10) actually strongly suggest not (just) an epistemic but (also) an inferential reading (i.e. a type of evidentiality), witness the original English formulation of (6) by means of must. This reading is still imaginable but less obvious in (8) and (9), however. And it is not obvious at all in (7). So the question arises whether the inferential meaning when present is not rather due to the context: in (6) and (10)
On *certainly* and *zeker*

(And maybe also in (8) and (9)) the evidence leading to the epistemically modified conclusion is explicitly mentioned in the text; hence the inferential meaning is strongly invited, and might thus be no more than an implicature. This is an issue for further research.

As Table 1 above indicates, the epistemic use of these adverbs is – somewhat surprisingly – the least frequent one, both in English (leaving the pragmatic use aside) and in Dutch: it constitutes only a small proportion of their ‘post-manner’ uses. Moreover, as Table 2 shows, epistemic cases are frequently vulnerable to ambiguity (in English 52.6% and in Dutch 37.5% of the epistemic cases in a maximal count), predominantly with the strengthening use.

One last interesting observation is the occurrence of one instance of *zeker* in this reading in an interrogative, namely (11) (there are no comparable cases in the English data):

(11) Welk klimaat zal het zeker niet worden? (CGN: fv400042)

“What climate will it certainly [sic] not become?”

It is well known that epistemic adverbs normally cannot occur in non-declarative contexts (e.g. Bellert 1977). The reason is that they can only be used ‘performatively’, i.e. to express the speaker’s own evaluation of the state of affairs at the moment of speaking, but not ‘descriptively’, i.e. to report on another person’s or a hypothetical epistemic evaluation to which the speaker is not committed, which is the case in an interrogative (Nuyts 2001). Weak adverbs such as Dutch *misschien* “maybe” and English *perhaps* do occur in interrogatives, but then they do not express an epistemic evaluation of the state of affairs, but rather serve as a speech act modifier, turning the question into a tendentious one or making the request more polite (cf. van der Auwera 1983; Nuyts 2001). However, such a usage appears impossible with stronger epistemic adverbs. In fact, (11) is a highly marked case, in the sense that it is bound to a very specific type of context: it is a question formulated by a teacher in a geography class, aiming to check whether the students have understood the subject matter. Thus, it is actually a blend between a declarative – the answer in the mind of the teacher – and a question inviting that answer from the students, whereby only the critical information in the former is replaced by a wh-term, but the rest of the declarative utterance, including the modal adverb, is simply retained in the interrogative. (Hence the adverb in (11) is not a speech act modifier, but really does express epistemic modality, unlike the use of *misschien* and *perhaps* in questions.)

4.2 The scalar use

In the scalar use *certainly* and *zeker* situate an ‘element’ on a scale or in a range of related ‘elements’ given or implied in the context (i.e. typically what scalar particles do – cf. for instance *even*, *still*, etc.). The concept ‘element’ is to be taken broadly, and can involve objects, places, ‘times’, states of affairs, etc. Specifically, the element affected by the adverb is presented as being minimally the case, suggesting that other elements
higher on the scale or elsewhere in the range may possibly apply as well. Examples are (12) through (14) (the latter actually demonstrates that, pace Van der Wouden (2000:80), the collocation *zeker en vast* “certainly and surely” can also be used in this way).\footnote{Zeker en vast and *vast en zeker* (see Section 4.1) are regional variants of the same collocation: the latter occurs in Netherlandic Dutch, the former in Belgian Dutch.}

(12) Since nineteen eighty nine of course, rateable values disappeared. When the wonderful new system the poll tax came in. And so all new houses built since that date, have had no rateable value assessment. And so you can’t use that system. So all new houses, certainly in Severn Trent region, since that date have had water meters. (BNC: G4U 278)

(13) Volgens de Oostenrijkse TV zijn *zeker* tien inzittenden zwaargewond geraakt. (CGN: fn006129)

“According to Austrian TV at least ten passengers were seriously injured”.

(14) De meesten of toch de helft van de klassen waar ik kom zijn er toch *zeker* en vast drie gemiddeld per klas die voetballen. (CGN: fv400158)

“In most or at least half of the classes where I go there are at least three on average per class who play soccer”.

When combined with negation, this meaning turns into an indication that at least the element affected by the adverb and the negative particle does not apply, and that even other, implied elements may not apply either (that is to say, the negation extends to conceptually implied elements). This is illustrated in (15) and (16):

(15) A: Would you have shared that view with them at that time? That [reading] It is not considered appropriate that Skelton should expand any further?  
B: I can’t answer that, because there’s *certainly* nothing in the text of the plan to suggest why it was not considered appropriate, and I don’t know I can’t instantly think what factors might have gone into that consideration. (BNC: FMP 520)

(16) Nee gij pakt niet en dan *zeker* niet als dat met ons samen is dan pakt de gij dus niet de Seppe mee hè. (CGN: fv900059)

“No you don’t bring along, and certainly/especially not when that would be together with us, then you don’t bring along Seppe, do you understand?”

As the examples show, this meaning can involve a strictly ‘numerical’ scale, as in (13) and (14), but more often it involves a more loosely or vaguely organized set or range of elements. In (12), for example, it involves possible places where one can find water meters in new houses, this being the case at least in the Severn Trent region, and
On *certainly* and *zeker* 

possibly/probably in other places as well (the instances in (15) and (16) are also of this type).¹⁴

English and Dutch differ somewhat in the syntactic possibilities of the adverb in this usage. First, as examples (13) and (14) demonstrate, in Dutch it can be used at the term level, i.e. as an adverb within a noun phrase, in which case it always involves a numerical scale. We have not found any such term-level uses in the English data, and intuitively they seem impossible. This does not mean that English lacks the numerical scale use of the adverb: this use also occurs with the sentence level adverb, both in Dutch and in English – (17) is an instance (*certainly* concerns the question whether the picture is worth a nine or a ten).

(17) So we’ve got lovely soft lighting throughout. Brighter on the side which is furthest away from us, and the fact that he’s, he or she, is eating a, you know a piece of fruit or something is quite appealing. I like that one. Not quite a ten because I think it’s you know, this hand here perhaps a little bit obtrusive but *certainly* I think that one’s worth a nine. (BNC: HM2 749)

Second, in all sentence-level uses in the Dutch data, scalar *zeker* affects one constituent only, rather than the complete expression, and in terms of word ordering it appears immediately before the constituent at stake. Thus, in (16) it only affects, and appears immediately before, the conditional subclause; another example is (18), in which it only affects the ‘spatial’ prepositional phrase and again appears immediately before it.¹⁵

(18) ‘t Vak Engels en ‘t vak Nederlands het geven is *zeker* op ‘t niveau waarop ik werk is heel verschillend. (CGN: fn000088)

“The English course and the Dutch course, teaching them, *certainly* at the level at which I work, is quite different”.

This kind of use is to be expected since usually the ‘elements’ in the scale or range are objects, or places, or subaspects (expressed in a nominal group or subclause) which are part of the state of affairs (expressed in the utterance) as a whole. English scalar

---

¹⁴. Van der Wouden (2000) restricts his ‘scalar’ category to the numerical type only – the other type would rather seem to correspond to his ‘in elk geval’ “in any case” modal particle category. But again, we see no reason to separate these two ‘types’ since they essentially involve the same basic meaning, the numerical type being no more than a special case of the general category (the difference between the types is only a matter of degree of strictness of the scale involved).

¹⁵. This type of syntactic use also affects the ‘syntactic behavior’ of the negation marker if present. As can be seen in (16), e.g., if the negation marker appears in the part of the clause preceding the adverb and the constituent in its scope, it has to be repeated immediately after the adverb. If the negation marker is only introduced in the clause somewhere after the cluster of the adverb and the constituent in its scope, however, there is no negation marker immediately after the adverb – compare the English translation in (16) to: *certainly especially when we are there, you don’t bring along Seppe, do you understand?*
certainly is also used in this way (see (12) and (15)). But in some instances, although the adverb still affects just one constituent, it does not appear immediately next to that constituent but elsewhere in the clause, as in (17) above, where it only affects the value nine. And in quite a number of instances it affects the entire clause, whereby the state of affairs expressed in that clause is then part of a set of states of affairs implied or mentioned in separate sentences in the wider context, as is shown in (19) and (20):

(19) Now I personally believe that most policy makers in Britain, and indeed in most other industrial countries, don’t really believe the Third World matters to them at all. You see our most recent budget was largely constructed even neglecting what was happening in other industrial countries. It certainly gave virtually no attention to the impact of issues from developing countries (BNC: KRG 873)

(20) What we’ve got is a tremendous gap in what people need to know in society, in all areas of society, to handle this computer revolution. So a big issue about whether how well we’re placed with regard to America and Japan and so on is how well placed are we to bring about this educational step, and I think we’re probably about as well placed as anybody else, certainly we haven’t made some of the mistakes that other people have made (BNC: KRF 228)

Why these syntactic differences between the two languages should exist is currently not entirely clear to us – it may have to do with the greater freedom in Dutch as compared to English in positioning adverbial elements in an utterance. Whether this explanation holds water or is sufficient is an issue for further investigation.

We have actually not found any traces of the scalar use of these adverbs in our diachronic sources. Apparently, this is a fairly ‘recent’ development, in both languages.

As Table 1 above shows, this usage is the most frequent one in Dutch, but in English it is only the second most frequent. Also, as Table 2 shows, in English, unlike in Dutch, this usage is characterized by a high number of ambiguous instances: 48.9% of all scalar cases in a maximal count, as compared to only 5.1% in Dutch. Moreover, the two ambiguous instances in Dutch, both involving ambiguity with an epistemic reading – among them (21) – do not even involve ‘real’ ambiguity for the hearer: in both there is insufficient information for the analyst (specifically, intonational information – the wider context does not help) to decide which of the two meanings was
meant by the speaker, but the ‘original’ hearer will probably not have had any problems in this regard. (In the epistemic reading, *zeker* in (21) would affect the entire clause, whereas in its scalar reading it would only affect *vanavond* “tonight”.)

(21) Maar ‘k ga ‘t wel *zeker* vanavond niet te laat maken. (CGN: fv701063)
    “But I am certainly not going let it get late tonight/but I am, at least tonight, not going to let it get late”.

In English, by contrast, alongside a few cases of ambiguity with an epistemic reading, this reading is very frequently ambiguous with the strengthening reading, and in most cases the ambiguity will most probably have been there for the ‘original’ hearer as well (see Section 4.3 below for examples). It is possible that this difference between Dutch and English correlates with the syntactic differences discussed above.

4.3 The strengthening use

In their strengthening use, *certainly* and *zeker* serve to reinforce an assessment of some kind (an opinion, attitude, evaluation – the notion has to be applied broadly) of a state of affairs, such as a moral, epistemic, aesthetic or quality judgment, expressed in the same or a preceding utterance. This often involves an assessment by the speaker himself, as in (22) and (23), in which it is contained in the same utterance as the adverb, or in (24), in which it occurs in the previous utterance by the same speaker. But this can also involve an assessment expressed by the interlocutor in a preceding utterance, as in (25) to (27). ((25) and (26) involve another instance in which the adverbs in the Dutch and English versions of Harry Potter match.) (25) and (26) are ‘strengthening’ in that the adverb in them underscores the ‘appropriateness’ of the judgment inherent in the interlocutor’s request/order that the speaker should perform the action at stake (see below).

(22) You provided yourself with the uniform before you came as a student nurse so that was alright for the first year, so you’d no uniform to provide. But I often wonder what the nurses today would think if that’s all they got. I know money did go much farther in those days but there *certainly* wasn’t very much to spare at all. (BNC: H4C 338)

(23) Dus dat is iets wat dat ik *zeker* nu moet nakijken want als er dus … (CGN: fv400157)
    “So this is something that I certainly have to check now because if…”.

(24) Most people who claim invalidity benefit I mean don’t need it anyway they’re people who’ve got, you know, the existing bad backs, people like that and that’s, I actually feel very strongly about that because I wouldn’t mind paying more money if the people who needed it got it. I *certainly* wouldn’t. (BNC: K6Y 539)
(25) ‘Right,’ said Professor McGonagall crisply, also moving to the door. ‘I’ll leave you to deal with Potter and Weasley, shall I?’ ‘Certainly,’ said Dumbledore. (PE: 354–355)

(26) ‘Goed,’ zei professor Anderling kordaat en ze liep ook naar de deur. ‘Rekent u dan af met Wemel en Potter?’ ‘Ja zeker,’ zei Perkamentus. (PD: 246)

(27) A: Wel hele leuke ouwe dia’s gezien en zo. B: […] Ja en val je daar niet bij in slaap? A: D’r werd veel gelachen om die dia’s. En dat hield je wel wakker. En d’r liep ook nog een hondje rond die dat allemaal wel heel gezellig vond al die mensen die op de grond moesten zitten. B: Oh die houdt je ook wel wakker natuurlijk. A: Ja die houdt je zeker wakker. (CGN: fn008037)

“A: I did see a lot of nice old slides and stuff. B: And didn’t you fall asleep then? A: There was a lot of laughter about those slides. And that kept one awake. And there was also a little dog running around, who found all that very nice, all those people that had to sit on the ground. B: Something like that also keeps you awake, of course. A: Yeah, it certainly keeps you awake”.

As indicated, this usage can also involve the strengthening of an epistemic judgment – but this is not the same as expressing an epistemic judgment. In fact, instances such as (28) and (29) allow both readings: either these mean that the speaker ‘considers it 100% certain that it is true that so and so’ (i.e. an epistemic use), or, more likely, they mean that the speaker ‘strongly ascertains that it is true that so and so’ (i.e. a strengthening use).

(28) It’s certainly true that government guidance is far more guarded in its support for new settlements […] but I would emphasize the dates of this guidance… (BNC: HVK 154)

(29) A: ‘t Was bij ons vooral in het eerste jaar dat je even een achterstand oploopt maar die heb je zo ingehaald bij zoiets als celbiologie maar voor de rest valt ‘t heel erg mee. B: Ja dat is zo. Ja dat is waar. Dat is waar dat is zeker waar ja. (CGN: fn000188)

“A: In our case it was particularly in the first year that you fall behind for a while but you catch up in no time for something like cell biology but otherwise things aren’t all that bad. B: Yes that is so. Yes that is true. That is true that is certainly true yes”.

18. Especially example (28) is an instance in which the adverb is used concessively. There are actually quite many concessive instances, both in the English and the Dutch data, involving (also unequivocally) epistemic, strengthening and scalar uses of the adverb. Van der Wouden (2000:77–78) classifies concessive cases into a separate category. There is no reason to do so, however; in such cases, the adverb simply retains its ‘original’ value (epistemic, strengthening or scalar, or ambiguous between two of these), and the concessivity is entirely due to interaction with the context.
As (25) and (26) illustrate, this meaning also covers absolutive uses of the adverb (see Section 3). In fact, all our absolutive instances can be taken to involve strengthening uses. A few of them are, admittedly, ambiguous with an epistemic reading, including (30) and (31): in these the adverb can either be understood as strengthening the opinion implicit in the tendentious question by the interlocutor (in (31) the tendency of the question is obvious, in (30) it is milder but present: ‘since medicine is more expensive in the US, the diagnostic system will be used there more often’); or it can be understood as involving an epistemic assessment of the hypothesis in the question (i.e. as meaning ‘it is 100% certain that what you ask me is the case’).

(30) A: Is it used much more in the United States where medicine is much more expensive altogether? B: Yes, certainly, the place where such diagnostic systems came from, I think, was certainly the States. (BNC: KRH 3348)

(31) A: Stel je voor dat jouw vader niet gestorven was dan had ze waarschijnlijk heel d’r energie nog veel meer op ’t maatschappelijk vlak gelegd want zo maatschappelijk was jullie man hè? Die wou ook iets met d’r intellect doen hè? B: Ja zeer zeker. (CGN: fn000776)

“A: Imagine your father had not died, then she would probably have put all her energy even more into social matters, because your man was so social, right? He wanted to do something with her intelligence too, right? B: Yes, certainly”.

In instances such as (25)–(26) above, or (32) and (33), however, an epistemic reading seems to be ruled out altogether. Understanding the adverb in (32), for instance, as meaning ‘I [the speaker] estimate the chances that I can come and hold this door to be 100%’ hardly makes sense. The only reasonable interpretation is a strengthening one: ‘I underscore your assessment that I should come and help you’.

(32) A: Andrew could you come and hold this door? B: Certainly. (BNC: F7G 515)

(33) A: Now can you take the documents? B: Certainly. . . (BNC: FUT 575)

The examples so far may suggest that this usage is basically identical in English and Dutch (cf. the fact that in all of them the adverbs can be used as mutual translations). But this is not entirely true: there are quite a few English but no Dutch cases in which what is being strengthened is highly implicit and often quite vague and hard to pinpoint. Usually, in such cases one would not – or even cannot – use zeker in Dutch. Examples are (34) through (36). Clearly, these are not epistemic uses: the meaning ‘I estimate the chances that the state of affairs applies to be 100%’ does not make sense in any of these. Rather, the meaning can best be characterized as something like ‘I firmly assure you that . . .’. As such, it is still a strengthenener, but then not so much of some kind of assessment expressed in the utterance (not even in (34): the adverb does not really strengthen the quality judgment in looks better as such), but rather of the ‘global stance’ behind the expression as a whole.

(34) A: Is there anything any good on? B: Only the Arabian Nights at two and that’s gone. A: Well I don’t like those things. B: And then there’s Clint Eastwood at
twenty to ten. A: Nothing in between? B: No. A: Well, the garden certainly looks better for your little clear up B: Well it looks A: doesn’t it? B: looks tidier doesn’t it? (BNC: KBB 10706)

(35) A: It don’t seem to be any good at all, whatsoever. So I believe it is rather a waste of money, it’s just talking shop, shuffling papers about, and we regret that in the time when resources are difficult, staffing could be put elsewhere instead of this B: Right, I note your comments. But don’t agree with them, but were certainly noted. (BNC: J9D 354)

(36) A: We maintain three point centres during the year don’t we because Sherburn is open. B: That’s right, Keith said that he certainly would take that into consideration. I said I’d let him know the total number of contact days or whatever with Sherburn over the period (BNC: H5E 384)

Our diachronic sources for English offer firm evidence for the existence of the strengthening use of certainly at least from 1300 onwards, and possibly even before that. One example (taken from Kurath 1959) is (37). Our sources for Dutch are less clear, but instances such as (38) (from Verwijs & Verdam 1885–1929) and (39) (from De Vries & Te Winkel 1864–1998) would seem to come very close to this usage.

(37) And der yee suer...Pat yes herd and sa gh al þis? Ya, certainl¦, pat soth it es We tak drightin til vr witnes. (Cursor Mundi, before 1325)

(38) Sekerlijc het ware sere pijnlijc te gelovene ... (Roman van Lancelot, 13th century)

“Certainly it would be very embarassing to believe ..”

(39) ’t Ghemeyne Volck ... , Ten maackt gheen onderscheet in een pompeuse ziel [...]: Zekers ick bedruyf may dat hier het Volck so bot is (Bredero, Spaanschen Brabander, 1617).

“The Mob, It does not differentiate in a [i.e. its] pompous mind: Certainly I regret that People are so dull here”.

As Table 1 above shows, this usage is, by a considerable distance, absolutely dominant for present day English certainly. For Dutch zeker it depends on whether one considers the minimal or the maximal frequencies: in the latter, the difference with the scalar use is marginal, in the former it is second in line. Table 2 indicates that this use is moderately vulnerable to ambiguity. In Dutch 18.4% of the strengthening cases in a maximal count are ambiguous, with one exception all of them with an epistemic reading – this is much as compared to the scalar use, but not much as compared to the epistemic use. In English ambiguity amounts to 38.4% of the strengthening cases in a maximal count, i.e. a relatively small number as compared to the two other uses discussed above. What is striking about English, however, is the very high number of

19. Thus it is surprising that this meaning seems to be lacking altogether in van der Wouden’s (2000) analysis.
On certainly and zeker

instances (in absolute counts) of ambiguity with the scalar reading. (There are also a few instances of ambiguity with an epistemic reading, comparable to such cases in Dutch – see above for examples.) Examples are (40) and (41):

(40) A: What impact do you think it has when you have a complex in which there is such a high level of unemployment in which so few people are working, [...] do you think it has certain implications or not? B: I certainly feel myself that it intensifies the individual problem. (BNC: GYK 145)

(41) So it's easy for me to sit here and say, we think we're doing a good job on the environment, because we're not threatened, we do have to find ways round existing chemicals, we may have to pay a higher price, but we're not threatened in terms of the core of our bottom line. I think companies certainly are interested in creating the right impression, in fact in creating the right environment, and they are willing, of course, to pay a price, but the argument which you put is if you threaten the companies existence, right. (BNC: JNL 163)

The causes for this type of ambiguity in English (but not in Dutch) are no doubt the syntactically often less 'transparent' way (in view of its semantic scope) of using scalar certainly (cf. the discussion of examples such as (17), (19) and (20) in Section 4.2), as well as the semantically more vague type of strengthening use of the kind in (34)–(36) above, neither of which seems to exist in Dutch. This often triggers uncertainty – probably not just for the analyst, but also for the 'original hearer' – as to which of these meanings is involved.

Finally, we should mention one remarkable use of strengthening zeker in our data. Like the epistemic use, intuitively the strengthening use would not seem possible in an interrogative, yet (42) does feature such a use.

(42) Zeg maar ga je zeker de vraag stellen aan ’t VVSG waar dat de verdere opvolging gebeurt van . . . (CGN: fv700038)

"Listen are you certainly [sic] going to ask the question to VVSG where the further follow up will happen of .. ."

As in example (11) in Section 4.1, this is no doubt a syntactic blend, in this case of a declarative containing a deontic assessment and zeker as a strengthener, i.e. 'you certainly must ask the question . . .', and an interrogative serving as a request to the hearer to ask the question.

4.4 The pragmatic use

The last type of use, occurring in our Dutch data, is what we will call a 'pragmatic' one – 'pragmatic' because the adverb does not express a semantic (strengthening of a) modification of a state of affairs, but operates as a speech act modifier: it turns a declarative into a special type of interrogative (one might call it a 'declarogative', in the vein
of Sadock 1971), viz. one in which the speaker asks for confirmation of the claim formulated in the declarative. Typical examples are (43) and (44); (45) illustrates how this usage can be employed for ironical purposes.

(43) Hij zal wel weten wat ie kan zeker? (CGN: fv400505)
    “He’ll know what he can do, won’t he?”

(44) ’Tijdens het eten moeten we proberen onze gasten een paar mooie complimenten te maken. Had jij ideeën, Petunia? ‘Ik hoor van Herman dat u een geweldige golfer bent, meeneer Bouwmeester… Die jurk komt zeker uit Parijs, mevrouw Bouwmeester?’ (PD: 8–9)
    “Now, we should aim to gel in a few good compliments at dinner. Petunia, any ideas? ‘Vernon tells me you’re a wonderful golfer, Mr. Mason… Do tell me where you bought your dress, Mrs. Mason.’”

(45) ‘Bang?’ mompelde Malfidus zacht, zodat Smalhart hem niet kon horen. ‘Voor jou zeker?’ zei Harry uit zijn mondhoek. (PD: 144)
    “Scared?” muttered Malfoy, so that Lockhart couldn’t hear him. ‘You wish,’ said Harry out of the corner of his mouth.

In this usage the adverb has a slight preference for clause-final position, as in (43) and (45) (i.e., comparable to a tag question), over a position within the clause, as in (44) (the former occurs in 55.2% of the minimal and 54.8% of the maximal number of pragmatic cases). The clause normally maintains declarative syntax, as in the above examples. The only exception in our data is (46). To a native speaker this is a weird construction, however, and it is probably a syntactic blend again: the speaker has an assumption in his/her mind, for which (s)he wants to seek confirmation from the hearer, which (s)he then considers formulating as a tendentious ‘zeker-declarogative’ yet ultimately renders in the syntactic shape of a yes/no question.

(46) Schrok Riet daar zeker wel van? (CGN: fn000946)
    “Was Riet surely [sic] surprised by that?”

Three cases of pragmatic zeker in our data are somewhat special – (47) and (48) are two of them. These have all the properties of the pragmatic instances of the kind in (43)–(45), but in addition zeker serves here to express the speaker’s feelings about the state of affairs, i.e. whether or not (s)he hopes that the state of affairs will apply. The expression is then clearly meant to influence the addressee’s actions.

(47) Je gaat toch niet overmorgen opnieuw naar de vrederechter stappen zeker in verband met ‘t één of ‘t ander hè? (CGN: fv900034)
    “You are hopefully not going to go to court again in a couple of days for one or another reason, are you?”

(48) Professor Anderling zei dat op 1 juni de eerste examens zouden plaatsvinden, over precies een week. ‘Examens?’ brulde Simon Filister. ‘We krijgen toch zeker geen examens?’ (PD: 211)
“Ten minutes into the class, Professor McGonagall told them that their exams would start on the first of June, one week from today. ‘Exams?’ howled Seamus Finnigan. ‘We’re still getting exams?’”

These instances all contain negation and toch (a modal particle which cannot be readily translated into English, and the meaning of which is very difficult to define, but at least in the present cases appears to pertain to the marking of a contrast between the present speaker and the interlocutor). However, the negation is clearly not necessary nor sufficient to get this reading: there are very many ‘normal’ pragmatic cases with negation, and one can perfectly well imagine cases without negation that do get this special interpretation, for instance (49):

(49) Je gaat toch zeker met ons mee?
“You are hopefully going with us, aren’t you?”

The presence of toch does seem necessary, but it is still not sufficient, witness (50), which takes the normal pragmatic reading of the kind in (43)–(45) (there are more examples of this in the data).

(50) Daar zijn toch zeker bekwaame mensen bij. (CGN: fv600215)
“There are surely knowledgeable people among them, aren’t there?”

What triggers this special reading, then, is a matter for further investigation.

Our diachronic sources for Dutch reveal instances of this use from the 19th century onwards, including (51) (from De Vries & Te Winkel 1864–1998) – but we did not find any traces of it in older stages of the language.

(51) Zou men niet zeggen dat gy spyt hebt omdat de schelm Hensmans op het schavot heeft gestaen? Ze moesten zeker dien nachtdief laten loopen, niet waer? (H. Conscience, Lambrecht Hensmans, 1847)
“Wouldn’t one say that you feel sorry because the crook Hensmans stood on the scaffold? They surely should have let that ‘thief in the night’ go, shouldn’t they?”

As can be seen in Table 1 above, the pragmatic use is manifestly present in our Dutch data, and as Table 2 shows, it is a very ‘recognizable’ use since it is hardly subject to ambiguity (only 6.5% of the maximal number of pragmatic instances).

For English certainly, however, the situation is quite different: we found only two instances which might tentatively be interpreted as pragmatic, (52) and (53). Especially the latter would at first sight seem to come fairly close to it (the Dutch equivalent using pragmatic zeker would then be: maar meneer X had toch zeker wel het recht om dat naar Y te sturen, of niet?).

(52) A: You came here in nineteen eighty, did you? B: Nineteen eighty, I’ve been here just on four years now. A: You certainly keep yourself busy, don’t you? B: Yes. (BNC: K64 586)
(53) **A:** If I had a vote that day, I certainly voted in favour of that, and so did members of all groups. Some members voted against. But *certainly* Mr. X had the right through standing orders to ensure that that was taken to county council, wouldn’t he? **B:** Yes sir, county council could make that decision. (BNC: JJJ 461)

In both instances the adverb can clearly also be interpreted as a strengthening case, however (of the ‘vague type’ exemplified in (34)–(36) in Section 4.3). Moreover, in all pragmatic instances of *zeker* in the Dutch data, translation by means of *certainly* appears completely impossible – in fact, none of the 9 instances in the Dutch version of Potter corresponds to *certainly* in the English version (in the English original one typically finds speech act verbs, as in (44) above, or mental state verbs such as *I suppose* or *I bet*, or *surely*, or tag questions). So it looks like this usage is at best marginal and probably entirely absent for English *certainly*. (There are also no traces of it in our diachronic sources.) Note, by the way, that this observation does not apply to *certainly*’s ‘sister adverb’ *surely*: the latter clearly does have the same pragmatic use as Dutch *zeker*.

5. Conclusions

In summary, *pace* our original impression from the Potter data, it turns out that English *certainly* and Dutch *zeker* behave in fairly comparable ways. The most important differences are the (near) absence of the pragmatic use in English, a use well developed in Dutch, and the existence of a weaker and a stronger epistemic use in Dutch, while English only features the stronger one. Otherwise, there are only smaller differences within the shared uses, viz. in the syntactic possibilities of the scalar use, and in the ‘semantic generality’ of the strengthening use, English allowing more vagueness in what is being strengthened than Dutch. Also comparable is the frequency of the shared uses in the two languages: the strengthening use is absolutely dominant in English (maybe as a ‘by-product’ of the higher generality of this meaning in English), while the scalar one is (slightly) more important in Dutch, but in both languages the epistemic use is clearly the minor one. One further difference between the languages – visible in Table 2 in Section 4 – concerns the vulnerability of the adverb to ambiguity: only 8.2% of the instances of *zeker* have it, but for *certainly* this amounts to 28.4%. This fact is no doubt to a large extent due to the differences particularly in the scalar and the strengthening uses. And it probably correlates with the general impression one gets when looking

20. A remarkable observation is that in English and Dutch alike, the weak epistemic adverb which can be used as a speech act modifier – *misschien* and *perhaps*, see Section 4.1 – can only modify an interrogative, the strong one only a declarative. In both cases, the result is a tendentious question – a more moderate one in the former, a more outspoken one in the latter.
at the data, that \textit{certainly} not infrequently, and certainly much more often than \textit{zeker}, seems to be used as a rather functionless stopgap.

Let us come full circle and briefly return to the diachrony of these adverbs (cf. Section 2):

i. Since the two variants of the epistemic use of Dutch \textit{zeker} appear to be present from the earliest sources onwards, it is difficult to estimate which of them is older: they could both have originated in the manner use of the adverb, or one could have developed out of the other.

ii. The pragmatic use in Dutch is likely to have developed out of the weaker epistemic use, possibly sometime in the 19th century, if we can trust our diachronic sources.\footnote{Their trustworthiness for most non-manner and non-epistemic readings is not very high though: except for the strengthening use in English, none of them is mentioned as a separate category in the different historical dictionaries consulted (we have only found traces by interpreting instances mentioned under other categories), hence it is not unlikely that the dictionary makers – since they have not specifically looked for these meanings/uses – have overlooked instances of them in their older materials.} This is in line with the intuition that this pragmatic use relates to a slightly lower degree of speaker certainty regarding the state of affairs: why else would one seek confirmation from the hearer? (consider, too, the fact that in the Potter data the corresponding English forms are systematically not of the strongest type). It also corresponds to the observation that in English \textit{surely} – which is overall somewhat weaker than \textit{certainly} – has also developed this pragmatic use. And it explains why English \textit{certainly} has not developed this pragmatic use: it does not have the weaker epistemic reading. This would then involve an instance of a diachronic development from an epistemic adverb into a kind of discourse particle, probably not unlike other developments of this kind described by Traugott and Dasher (2002:152ff.).

iii. The scalar use is clearly related to the stronger epistemic meaning of the adverb, in both languages: in a way, it still reflects the meaning that ‘there is certainty that the element in the scope of the adverb applies’ – but it does much more than that, of course. Our diachronic sources suggest that this is a very recent development; nevertheless, the new meaning is already well established, as is suggested by the very small number of ambiguities between the original epistemic and the derived scalar meaning.\footnote{For evidence that frequency of ambiguities can be used as a diagnostic tool in meaning development, see Van Ostaeyen and Nuyts (2004).}

iv. The strengthening use of \textit{certainly} and \textit{zeker} is also clearly related to the stronger epistemic meaning. It may have emerged – in or even before Middle English and (probably) Middle Dutch – as a conversational implicature when the adverb was used in a context in which the issue of the existence of the state of affairs had already been brought up before. The reinforcing effect arising from the repeti-
tion of the epistemic statement may subsequently have become an autonomous use, which has then been generalized to all circumstances in which ‘opinions’ (of whatever kind) are at stake.

What do all these observations mean in terms of the basic issues mentioned in the introduction? It is striking to see how adverbial markers of epistemic certainty (in both languages – but it is probably not outrageous to assume wider validity for this observation) develop a fairly wide range of meanings/uses beyond the domain of ‘qualifications of states of affairs’ or ‘tense-aspect-modality categories’, uses which then also become much more frequent than the epistemic qualifying use. This is much less the case for markers of epistemic possibility (in Dutch and English these have only developed the pragmatic speech act use, which is, moreover, not very frequent) and hardly at all for adverbs expressing probability (cf. Nuyts 2001). This is no doubt related to the conceptual status of these different epistemic values. Certainty is obviously the ‘default’ status, and definitely the desired status, for a state of affairs in one’s conceptualization of the world, while uncertainty (of whatever degree) is a marked status, and one which causes ‘conceptual instability’ (Nuyts 2001:364). This probably also means, then, that talking about certainty has a quite different conversational role from talking about uncertainty, at least in the sense that conceptual uncertainty will typically invite conversation about the status of the state of affairs (in order to achieve a higher degree of certainty about it), but certainty much less so.\footnote{For circumstances under which certainty does invite conversation, see Simon-Vandenbergen (this volume).} Maybe this also means, then, that terms specifically dedicated to expressing epistemic certainty are relatively less often ‘needed’. At the same time, however, certainty is, unlike weaker epistemic values, no doubt quite salient as a conceptual notion \textit{per se}, hence it may easily be taken as an ‘example’, or a ‘metaphor’, for ‘strength’ in other areas. The combination of these two facts may then cause expressions for epistemic certainty to be employed for other purposes much more easily. This is all quite speculative, of course, and further in-depth research (and theorizing) will have to reveal whether this line of thought leads anywhere.

References

On certainly and zeker


Prenominal possessives in English
Function and use*

Evelien Keizer
University of Amsterdam

This paper examines the circumstances favouring the use of a prenominal possessive in English, even if, syntactically and semantically, a postnominal construction would also have been acceptable. An evaluation of earlier treatments of the difference between the constructions reveals two approaches. The traditional approach typically yields multifactor accounts, concentrating on the constraints on both constructions, while the theoretical approach is characterized by a tendency to reduce the difference to one underlying principle. Using authentic examples, this paper proposes a theoretically-based multifactor account which adds a major cognitive-pragmatic factor to those traditionally distinguished, and which tries (a) to resolve the question of whether any factor(s) is (are) more prominent, and (b) to explain, at least for some factors, why it is that they favour a particular construction.

1. Introduction

English has a number of ways in which to express possession within the noun phrase, the most important being the prenominal possessive (or genitive; e.g. the author’s opinion) and the postnominal possessive (or of-possessive; e.g. the opinion of the author). In this paper I will examine the circumstances that favour the use of a prenominal possessive, even if, from a syntactic and semantic point of view, a postnominal construction would also have been acceptable. Both types of construction have, of course,

* I would like to thank the editors of this volume for their useful comments on an earlier version of this paper. Any remaining mistakes or shortcomings are, of course, entirely my own.

1. There are of course other ways of expressing possession within the noun phrase, such as the postnominal genitive (e.g. a friend of Peter’s) and, possibly, the use of certain adjectives (e.g. Dutch cities); these, however, will not be included in the discussion here.
been described in considerable detail both in traditional grammars and in more theoretical frameworks, whereby the former have been first and foremost concerned with specifying the circumstances which disallow either of the two types of possessive, whereas the latter are primarily interested in the internal structure of the NPs in question. Much less attention has, however, been paid to the question of what determines a speaker's choice when both types seem to be syntactically and semantically equally acceptable.

There are a number of exceptions to this general tendency. Thus, R. Hawkins (1981), J. Hawkins (1994), Taylor (1996) and Rosenbach (2002) all do address the question of what determines the choice between a prenominal and postnominal possessive. The first three accounts try to do so in terms of one basic underlying principle (animacy, complexity and topicality, respectively), while Rosenbach (2002) looks at the interaction between three factors (animacy, topicality and the possessive relation). As we will see, however, these accounts all have their weaknesses. The present paper, although limited in scope, will be an attempt to remedy some of these shortcomings by considering the role of a variety of factors as well as the relation between them. Particular attention will be given to the cognitive element and the role of context, and authentic examples will be used to support the proposal. Such an approach, it is felt, will not only lead to a better understanding of the use of possessive constructions in English, but may ultimately also help to obtain a clearer picture of some of the more general cognitive processes underlying linguistic expression.

In what follows I will start by outlining the various factors that may influence a speaker’s choice for a particular expression. First, Section 2 will list the factors commonly identified in traditional accounts of the English genitive. In Section 3, I will briefly discuss a number of earlier proposals. Section 3.1 will be devoted to three proposals, by R. Hawkins (1981), J. Hawkins (1994) and Taylor (1996), which try to reduce the difference between the two possessive constructions to one basic underlying principle; Section 3.2 will discuss Rosenbach’s (2002) interactive account. Next, in Section 4, I will propose a more comprehensive multifunctional account, which will be applied to a number of authentic examples of possessive constructions. Eventually, I hope to be able not only to add a major (cognitive-pragmatic) factor to those traditionally distinguished, but in addition to determine whether any – and, if so, which – factor is more prominent (or basic) than the other factors, and to explain, at least for some factors, why it is that they favour a particular construction. Unfortunately, the scope of this paper does not allow for detailed discussions of both prenominal and postnominal constructions. The emphasis in this section will therefore be on prenominal constructions, although a brief section on postnominal constructions has been included for the sake of comparison (Section 4.3). Section 5 will present some conclusions.

Finally, I would like to add a brief note on the examples and terminology used. Most of the discussion will be based on authentic language material, taken from the ICE-GB Corpus, a fully parsed, one-million-word corpus of spoken and written English composed at the Survey of English Usage, University College London (for a detailed description, see Nelson, Wallis, & Aarts 2002). As for the terminology used, I will
refer use the term prenominal possessive (construction) to refer to what is traditionally known as the genitive and the term postnominal possessive (construction) to refer to the of-variant. The two NPs contained within these constructions will be referred to as the possessee (i.e. the head noun) and the possessor (complement/modifier), even if the relation between the two concepts is not strictly speaking one of possession (but, for instance, a (body) part-whole, part-attribute, kinship or locative relation, or the relation between an event and a participant in that event, or, indeed, a host of other relations; see also Janssen, this volume).

2. Traditional accounts: Interacting principles

Generally speaking, traditional linguists assume that the use of prenominal or postnominal possessive is not subject to strict rules or constraints; instead, they claim, we are dealing with tendencies and preferences. Some of these are of a syntactic or semantic nature; these, in particular, have been described in much detail. Other factors are of a pragmatic or stylistic character and tend to be dealt with rather cursorily. Many readers will no doubt be familiar with these tendencies; however, as they form the basis of the discussion to follow, I will nevertheless mention them briefly.

Let us begin with the semantic and syntactic factors, as these seem to be most widely accepted:

I. Animacy
Among the noun classes most frequently found as possessor nouns in prenominal position are proper names referring to persons (Mary’s car), common nouns referring to persons (the boy’s bicycle) and nouns denoting higher animals (our dog’s teeth). In addition, genitive constructions are often used with collective nouns denoting groups of people (the government’s plans, the company’s future).

II. Length/complexity
Another factor which plays an important role in choosing one construction rather than the other is the (relative) lengths or degrees of complexity of the possessor and possessee. Complex and/or lengthy possessors tend to appear in postnominal position, while simple, short possessors are more likely to occur in prenominal position. The reverse is true for possessee nouns. In (1) for instance, where the possessor is short and structurally simple and the possessee long and complex, the prenominal construction is obviously preferred:

(1) a. I was delighted to get your letter, of course, and very pleased to hear about the Council’s increased interest in the Hospice Movement. <ICE-GB:W2F-004 #75:1>
b. On the European stage reassured perhaps by Germany’s confused reaction to the war and its continuing problems of unification France seems readier to play a constructive role <ICE-GB:S2B-039 #90:3:A>

III. Ambiguity

In some cases the choice between the two constructions is determined by the wish to avoid ambiguity. Such ambiguity is likely to arise in constructions containing some other prenominal or postnominal element which may be interpreted as modifying either the possessor or the possessee noun (e.g. a poor doctor’s daughter or the daughter of a doctor of considerable intelligence).

IV. Morpho-syntactic restrictions

According to Huddleston and Pullum (2002:478), there is a preference for the postnominal position (even with a human possessor) in those cases where, in spoken language, the possessor is potentially ambiguous between a singular and a plural referent (e.g. my parents’ house; cf. Kruisinga 1932; Poutsma 1914).

V. Other restrictions

Occasionally, the use of a genitive may be favoured because of other semantic properties of the construction. Consider, for instance, the use of the genitive in the examples in (2). Here, for the genitive NP to be interpreted as the agent (or first argument) of the verb which forms the base of the past participles owned and managed, it must appear in prenominal position.

(2) a. and we rely on the settlement by the BBC’s wholly owned subsidiary <ICE-GB:S2A-064 #45:2:A>  
   b. In 1995, the supply contract with GrandMet’s 1,540 managed pubs runs out. <ICE-GB:W2C-016 #113:3>

Other restrictions are of a (lexico-)semantic nature. Thus, as Huddleston and Pullum (2002:482) observe, the adjectival element own is “virtually restricted to occurring after a genitive subject-determiner”, even if an of-construction might have been preferred on other grounds.

In addition to these semantic and syntactic factors, some descriptive accounts also recognize the importance of certain pragmatic and stylistic considerations:

I. Centrality/prominence

There are also cases, however, where the speaker may prefer a construction for pragmatic reasons. Thus, according to Zandvoort (1958:113), the postnominal position gives more prominence to the possessor than the prenominal position (cf. Kruisinga

2. In the codes following examples from the ICE-GB Corpus, S and W stand for spoken and written, respectively.
1932; Poutsma 1914). These early considerations of prominence are clearly forerunners of the more recently introduced theoretical notions of Topicality and (End) Focus (e.g. Quirk et al. 1985).

II. Style and genre
Finally, some linguists note that occasionally the use of a construction is influenced by the style of a discourse (e.g. formal vs. informal language, rhetorical effect) or the genre (e.g. poetry vs. prose, spoken vs. written language) (e.g. Altenberg 1982; Jespersen 1933; Poutsma 1914).

It will be clear that a speaker’s use of a prenominal or postnominal possessive can be motivated in a number of ways. In most of these descriptive approaches, however, there is no indication of why in a particular case one particular factor should take precedence over all the other factors; nor is there any suggestion that some factors may have more weight than others. Moreover, no attempt is made to establish the independence of the various factors. Thus, there may well be a (causal) relation between two or more of these principles (e.g. topicality may be higher among animate entities; focal NPs may typically be longer or more complex than topical ones). As a result, the explanatory and predictive power of these accounts is limited, as it is almost always possible to account for the use of a particular construction by invoking at least one of the various factors, even if other conditions are obviously violated.

3. Some earlier accounts

3.1 The search for a single underlying principle

To remedy these shortcomings, a number of linguists have proposed theoretically based approaches which seek to reduce the difference between the prenominal and postnominal possessive to one basic, underlying principle. In this section I will briefly discuss and evaluate three accounts of this kind, each of which selects a different underlying principle.

3.1.1 R. Hawkins (1981): Animacy

One of the most widely accepted semantic principles influencing the choice between a prenominal and a postnominal possessor is the presence in the possessor noun of the features [+animate] and [+human]. This rather crude characterization has been further elaborated by R. Hawkins (1981), who regards the restrictions observed as resulting from a semantic strategy whereby human nouns have linear precedence over non-human nouns. According to this strategy, if one noun denotes a human being and the other a non-human entity, the construction in which the human noun precedes the non-human noun is more acceptable. Refinements of this principle include a separate position for nouns denoting body-parts (whether or not used to refer to
human body-parts) and a distinction between non-human animate and non-human inanimate. The complete hierarchy is given in (3); some examples are given in (4).

(3) [human] < [human attr] < [non-human animate] < [non-human inanimate]

(4) a. two human Ns: Mary’s brother/the brother of Mary
b. human N + human attr: Peter’s legs/the legs of Peter
c. human N + non-human anim N: Mary’s dog/the dog of Mary
d. human attr + inanimate N: the foot of the mountain/the mountain’s foot
e. non-human anim N + inanim N: the cat’s basket/the basket of the cat
f. two inanimate Ns: the ship’s funnel/the funnel of the ship

However, although this hierarchy seems to work well in a very general sense, it nevertheless leaves many things unaccounted for. To start with, it is by no means clear where abstract nouns would rank in the hierarchy. Presumably, they would fall under non-human inanimate (or even lower on the scale, as suggested by Mackenzie 1983:46), but their behaviour does not seem to warrant such classification. Thus we have *the basis of their friendship*, but hardly *their friendship’s basis*; the intensity of his emotion, but not *his emotion’s intensity*. Yet, this may not be altogether surprising, as there are, in fact, quite a number of concrete inanimate nouns as well which do not (always) behave as predicted by the hierarchy. Thus, although both expressions in (4f) are acceptable, this does not hold for any conceivable combination of two inanimate nouns: the front of the house is clearly preferred to *the house’s front*, while the reverse seems to be true for constructions like the room’s Persian carpet and *the Persian carpet of the room*, all of which should be equally acceptable according to the hierarchy given in (3). Moreover, the hierarchy does not explain the preference (in terms of frequency of occurrence) of Mary’s brother over the brother of Mary; nor does it explain what determines the choice between the two constructions where – semantically and syntactically – they seem equally acceptable; or even, occasionally, the choice for a non-preferred construction.

3.1.2 Hawkins (1994): complexity

J. Hawkins (1994) is an ambitious attempt to explain all word order phenomena, including the difference between prenominal and postnominal possessives, solely in terms of structural complexity. One of the basic notions in Hawkins’s theory is that of the Constituency Recognition Domain (or CRD), which consists of the minimal number of nodes needed for the addressee to recognize the structure of a relevant part of a tree:

The CRD for a phrasal mother node $M$ consists of the set of terminal and non-terminal nodes that must be parsed in order to recognize $M$ and all the ICs [= Immediate Constituents] of $M$, proceeding from the terminal node in the parse string that constructs the first IC on the left, to the terminal node that constructs
the last IC on the right, and including all intervening nodes that they construct. (Hawkins 1994: 58–59)

The CRD is seen as a psychologically significant subset of the total set of nodes in the tree, on the basis of which important parsing decisions can be made. In examples (5) and (6) this decision concerns the structure of the VP.

(5) a. I gave the valuable book that was extremely difficult to find to Mary
    b. S [NP VP [V NP [Det A N . . .]] PP [P NP]] (see Hawkins 1994: 59)

In (5), the CRD of the VP starts with the terminal node constructing the first IC on the left (V), i.e. gave (the only terminal node in V) and ends with to (the terminal node constructing the last the last IC on the right, i.e. the PP). The CRD for (6) is considerably less complex than that for (5), as here the CRD ends with the first terminal node of the final NP, the.

(6) a. I gave to Mary the valuable book that was extremely difficult to find

Hawkins’s overall approach is thus to explain word order variation in terms of the number of immediate constituents (ICs) within the CRD for a phase or clause relative to the number of non-immediate constituents or words (a construction’s IC-to-non-IC or IC-to-word ratio): the higher the ratio, the more optimal the word order.

Unfortunately, however, it turns out that Hawkins’s generalization cannot account for the position of ICs relative to the head (i.e. prehead or posthead position); here, the IC-to-word ratios either fail to favour the preferred (or indeed only possible) word order, or may even favour a non-preferred word order. For these cases, Hawkins therefore resorts to a different distinction: that between (possibly) single-word ICs and multi-word ICs. This leads to the following performance generalizations for the English NP (Hawkins’s 1994: 291):

(7) a. If an IC is prenominal within NP, then it is possibly single-word in the CRD for NP.
    b. If an IC is necessarily multi-word in the CRD for NP, then it is postnominal.
    c. If an IC is possibly single-word in the CRD for NP, then it precedes all necessarily multi-word ICs.

By way of illustration, consider Hawkins’s example of the NP yellow books. On Hawkins’s original approach, the NPs yellow books and *books yellow would have the same IC-to-non-IC ratio (2/2 in both cases), which of course fails to account for the ungrammaticality of the latter construction. The generalizations in (7) do not seem to do much better. It is true that they correctly predict the acceptability of yellow books, since the prenominal IC yellow is indeed a single-word IC. Note, however, that the generalizations in (7) do not exclude the possibility of a single-word IC appearing to the right of the head, thus accounting for the fact that there are certain adjectives that typ-
ically, or necessarily, appear in posthead position (available, present, proper), as well as constructions which only allow the adjective in postnominal position (something useful, anyone intelligent, somewhere quiet). Where both positions are available, this is also reflected by the respective IC-to-word ratios; thus, both books available and available books will have an optimal score (2/2). The same, however, is true for yellow books and *books yellow, as well as for the people present and *the present people (both 3/3 on Hawkins’s analysis). Thus neither the generalizations in (7), nor the IC-to-word ratios, can explain why in most cases only one order is allowed.

Observe further that the addition of an adverb, such as very, makes no difference on either approach. Thus, it does not change the IC-to-word ratio of the NP: since yellow, as the head of the AP, constructs the IC, the adverb falls outside the CRD for the NP (Hawkins 1994:289). In the generalizations in (7) the fact that a multi-word IC like very yellow can appear prenominally is captured by the phrase possibly single-word IC in (7a): very yellow may be a multi-word IC, but is still a possibly single-word IC.

Next, let us apply the generalizations in (7) to prenominal and postnominal possessives. According to Hawkins’s (1994:292) analysis of possessive constructions, both the prenominal possessor (e.g. the king’s) and the postnominal possessor (e.g. of the king) consist of a head category, Poss (i.e. -s) and P (of), respectively, plus an NP:3

\[
\text{(8) a.}
\]

\[
\text{NP}
\]

\[
\text{PossP}
\]

\[
\text{Poss}
\]

\[
\text{NP}
\]

\[
\text{the king}
\]

\[
\text{’s}
\]

\[
\text{daughter}
\]

\[
\text{b.}
\]

\[
\text{NP}
\]

\[
\text{Det}
\]

\[
\text{N}
\]

\[
\text{PP}
\]

\[
\text{P}
\]

\[
\text{NP}
\]

\[
\text{The}
\]

\[
\text{daughter}
\]

\[
\text{of}
\]

\[
\text{the king}
\]

Subsequently, Hawkins claims that the distribution of the two constructions is entirely consistent with (7), since the prenominal PossP can dominate just a single word, and frequently does so (e.g. John in John’s daughter), while the PP (in this case of John)

---

3. Hawkins does not give the exact analysis of the construction in (8b); nor does he provide – anywhere in his book – a general system for analysing clauses and phrases. The analysis given here (with Det, N and PP as daughters of NP) is therefore based on the (rather few) examples Hawkins gives.
is of necessity multi-word, and is exclusively postnominal (ibid.). However, if Poss (-'s) is a separate category, functioning as the head of a PossP, whether or not the prenominal PossP dominates a single word is immaterial, since the NP preceding -'s will not be part of the CRD for the top NP: after all, it is the suffix -'s, as the head of PossP, which constructs the IC, which means that the NP the king falls outside the CRD for the top NP (cf. the adverb very in Hawkins's analysis of very yellow books).

In other words, whether we have the king's daughter or John's daughter should make no (or little) difference, as in both cases the prenominal NP (the king and John) is not an immediate constituent of the top NP and thus falls outside the CRD for this NP. Nevertheless, Hawkins prefers to classify prenominal genitives as possibly single-word ICs; in doing so, however, he applies the single-word/multi-word distinction not to the NP as a whole, but to the NP within the PossP (in which case the generalization in (7a) does not apply).4

Hawkins's approach suffers from a more general weakness as well. Thus the generalizations in (7), as well as the notions of necessarily or possibly single- or multi-word IC they employ, merely describe the pattern observed (i.e. genitive NPs occur in prenominal position, of- phrases in postnominal position), but do not explain why one construction is used rather than the other where both are possible (equally optimal according to Hawkins's system). Nor do they account for the large number of exceptions to the optimal word order patterns identified by Hawkins.5

3.1.3 Taylor (1996): Topicality

Taylor (1996), finally, regards the difference between genitive and of-constructions as a reflection of the cognitive-pragmatic status of the two nouns. For Taylor the function of possessives can be characterized as follows:

The special character of the possessive construction lies in the fact that it invites the hearer to first evoke the possessor entity, and conveys that the possessee nominal is to be located in the neighbourhood of the possessor. The import of the possessive phrase is thus to make explicit the mental path that the hearer must follow in order to identify the target. (Taylor 1996: 17; original italics)

This approach subsequently serves “as a basis for differentiating the prenominal possessive from its close competitor in English (i.e. postnominal of) . . . “, as the different word orders found in the prenominal possessive and the postnominal of-expressions are taken to reflect the different mental routes the hearer is instructed to follow (Taylor 1996: 18). Consider the phrases the company’s director and the director of the company. According to Taylor, the prenominal possessive typically functions as a ‘reference point’

4. This is not to say that, even in Hawkins’s system, the complexity of possessor NP is immaterial: if the CRD for the NP within PossP or PP is non-optimal, this will reduce the overall score of the possessive construction.

5. For a more detailed discussion of Hawkins’s proposals, see Keizer (to appear).
Evelien Keizer

(e.g. Langacker 1991:170, 2002:327–329), instructing the hearer to first identify the (presumably) highly salient entity the company, and then to identify the referent of director with respect to that company. The prenominal possessive does not, however, encode any specific relation between possessor and possessee (see also Janssen, this volume). In postnominal possessives, on the other hand, the element of “explicitly elaborates the relational character of director (a ‘director’ is necessarily the director of something)”; moreover, definite reference is achieved not by virtue of the prenominal possessor, but in virtue of the initial the (Taylor 1996:18).

A second essential property of prenominal possessors, related to their function as referent points, is that they are topical (Taylor 1996:205ff.; cf. Mackenzie 1983:48). For Taylor (1996:207), the English possessive construction grammaticalizes the reference point function and can therefore “be regarded as a kind of ‘local topic’, which, like the Chinese-style topic, delimits the referential possibility of the possessee nominal.” A topic, in Taylor’s sense, has to do with the accessing of mental entities, or, using Chafe’s terminology, the activation of concepts:

An active concept is one that is currently lit up, a concept in a person’s focus of consciousness. A semi-active concept is one that is in a person’s peripheral consciousness, a concept of which a person has a background awareness, but which is not directly focused on. An inactive concept is one that is currently in a person’s long-term memory, neither focally, nor peripherally active. (Chafe 1987:25)

There are, of course, various ways in which a concept can be activated. In the first place, a concept may be evoked by the discourse context (verbal or non-verbal) in which it occurs. Concepts may, however, also be activated indirectly, entering a person’s focus of consciousness through activation by a “frame”, i.e. a body of knowledge conventionally associated with an already named concept, or with some feature of the situational context (Taylor 1996:212; Chafe 1987). In addition, there may be reason to assume that some concepts, by their very nature, are inherently more accessible than other concepts, irrespective, almost, of discourse context (“inherent” topicality; Taylor 1996:211). This would account for the fact that the general pattern is for a human possessor to serve as a reference point for an inanimate target, as human possessor are inherently topicworthy (cf. Comrie 1981:191; see also Mackenzie 1983:47–48).6

Once we accept that the basic requirement for the use of a prenominal possessive is that the possessor is topicworthy, Taylor (1996:212–213) continues, it will not come as a surprise that these possessors often refer to entities mentioned in recently preceding text, tend to be discourse topics, are overwhelmingly definite and often take the form of a pronoun. Possessee nouns, on the other hand, typically introduce new, previously unnamed entities into the discourse which are usually not referred to again; as such they are, pragmatically speaking, the exact opposite of the possessor:

---

6. A cognitive basis for this pattern is provided by Deane (1987:194–195), who suggests that inherent topicworthiness is a result of several interrelated psychological factors, such as egocentricity, basic level primacy (Rosch et al. 1976; Lakoff 1987; Taylor 2003) and acquisition.
The crucial point, with respect to referent continuity, is that *possessee nominals are maximally differentiated from possessors.* The possessor is a maximally topical nominal, which functions as a reference point for a maximally non-topical possessee. (Taylor 1996: 217–218; original italics)

To illustrate his point, Taylor (1996:213) quotes an example from Osselton (1988), given in (9). Here the abstract noun *dispute* appears as the genitive because this is what the text is “about”. Had it not been the topic of discussion, Osselton argues, an alternative wording – *the critical issue of the dispute* – would almost certainly have been preferred.

(9) More than 15,000 Ford workers were laid off last night . . . duty free perks, although the incentive, were not the dispute’s critical issue. (Osselton 1988:139)

Now, although I agree with Taylor that prenominal possessors are typically topical (in the sense of “activated”), I am sceptical about his claim that NPs in postnominal *of*-phrases are not. Instead, I believe that in both cases the possessor typically has an “anchoring” function, linking the intended referent to the hearer’s knowledge base (see Keizer 1992), a function most efficiently performed by a topical, identifiable possessor. What differs is the mental path the hearer is instructed to follow; not the topicality of the possessor NP. Thus, it is perfectly normal for postnominal possessors to be (semi-)active and to be (related to) the discourse topic. A case in point are postnominal possessors in the form of a proper name. Taylor’s assumption (based on an earlier small-scale study, Taylor 1991) is that proper names and pronouns typically occur in prenominal position, which is indeed what we may expect if the prenominal possessor serves as a reference point for the hearer. A search in ICE-GB, however, confirms this for pronouns, but not for proper names:7 proper names, both active and inactive, occur frequently in postnominal position.

Nor is it the case that all prenominal possessors are necessarily discourse topical. We have seen that, following Chafe, Taylor defines topicality in terms of activatedness: possessors typically denote active, or at least semi-active, concepts, while possessees denote new, unmentioned concepts. From Chafe’s (1987:25) original definition, however, we can only conclude that semi-active concepts are typically not discourse topical: they are in a person’s “peripheral consciousness” and are not “directly focussed on”. In other words, although I agree with Taylor that in prenominal possessive constructions the possessor and possessee tend to differ in terms of topicality and activatedness, I do not agree that it is the topicality/activatedness of a prenominal possessor which distinguishes it from the possessor in postnominal *of*-phrases.

7. A quick and dirty search shows 1916 prenominal occurrences of proper names and 2024 postnominal proper names (includes all possible semantic and syntactic relations represented by ‘s and of).
3.2 Interactive principles: Rosenbach (2002)

Rosenbach (2002) offers an interactive account of what determines a speaker’s choice between an ‘s-genitive and an of-genitive. On the basis of earlier treatments of possessive constructions (including Hawkins 1981 and Taylor 1996, 2003), she selects three factors that are likely to influence this choice; subsequently she carries out an experiment to test the predictions she has formulated concerning the role of and interaction between these factors.8

The three factors that are tested by Rosenbach are animacy, topicality and the possessive relation. The reason for selecting these three factors is that these together are regarded as determining the ‘naturalness’ of a possessive construction in a given discourse situation, whereby, following Dressler et al. (1987:11), naturalness is defined in terms of what is ‘easy for the human brain’ (see Rosenbach 2002:105). The concept of naturalness, in turn, is regarded as a bridge for an application of iconicity, which is also described in terms of ease of processing (Rosenbach 2002:105–106, 110–111). The factors animacy, topicality and possessive relation are taken to reflect two natural/iconic principles Rosenbach (2002:126–127). In the case of animacy and topicality, the underlying iconic principle is that of linear sequencing: [+animate] possessors are considered to be higher on the animacy hierarchy than [–animate] possessors; as such [+animate] possessors may be expected to precede the possessee, resulting in a preference for an ‘s-genitive, while [–animate] possessors are predicted to favour the of-genitive. Similarly, since [+topical] possessors are higher on the topicality hierarchy than [–topical] possessors, the former are expected to precede and the latter to follow the possessee. In the case of the possessive relation, the iconic principle reflected is that of conceptual distance, the prediction being that the closer – i.e. more prototypical – the possessive relation, the more likely the use of the ‘s-genitive, while the of-genitive is predicted to be preferred when the possessive relation is non-prototypical (see below).

To test her hypotheses, Rosenbach used an elicitation test, whereby 104 British English and American subjects were given short text passages from novels containing one or more possessive constructions, and at various points were asked to choose between an ‘s-genitive and a of-genitive. Subjects were male and female and represented various age groups. All subjects had a relatively advanced level of education (Rosenbach 2002:128).

The results were as Rosenbach predicted: [+animate], [+topical] possessors with a prototypical possessive relation with the possessee strongly favoured the use of an ‘s-genitive, whereas constructions with [–animate], [–topical] possessors and a non-prototypical possessive relation strongly favoured the use of an of-genitive (Rosenbach 2002:149). Examination of the other possible combinations of the six features showed that animacy was the most influential and the possessive relation the least influential.

---

8. Rosenbach looks at Modern English as well as earlier stages of English. What follows only concerns the synchronic part of Rosenbach’s study. For a multi-functional approach to the genitive/of-construction alternation in 17th-century English texts, see also Altenberg (1982).
Prenominal possessives in English

factor (Rosenbach 2002: 147–148). Rosenbach subsequently discusses the influence of the other variables, i.e. the differences between British English and American usage and the differences between different age groups (‘change-in-progress’).

Although the use of an experimental study is a great improvement on the introspective approach used in previous studies of the possessive, and although the theoretical notions used by Rosenbach to account for the patterns observed are definitely interesting, the experimental design unfortunately suffers from a number of weaknesses. First of all, there is, of course, the fact that only three factors were tested, whereas many more may be assumed to play a role. Given the importance of the notion of ‘ease of processing’ in explanation of her findings, it is particularly surprising that the complexity factor was not included in the experiment. This, however, was a deliberate choice, which Rosenbach (2002: 132) justifies as follows:

... rightbranching expansions of the possessor or the possessum may affect the choice of the genitive construction considerably. Therefore no items have been included in which the possessor or the possessum is postmodified.

Premodified possessors and possessees are excluded for the same reason. This is rather odd for a study intended to explain why speakers choose one construction rather than the other in particular circumstances. As a matter of fact, Rosenbach’s experimental design does not allow her to achieve this aim; it merely tests the relative influence of the three previously selected factors. As such little is added to the discussion, as there is no way of establishing new factors or of refining the factors chosen.

An additional disadvantage is that the operational definitions given of the three factors tested are (of necessity, no doubt) greatly oversimplified. Thus, the feature [+animate] is reserved for personal nouns only, while [–animate] is restricted to concrete nouns; thus all proper nouns, pronouns, collective nouns and abstract nouns are excluded from the experiment (Rosenbach 2002: 112–113). The feature [+topical] is assigned to second mention, definite possessors only, while [–topical] is defined as first mention and indefinite; as is well known, however, the relation between topicality and definiteness, as well as that between topicality and (textually) given is certainly not one-to-one. Furthermore, [+topicality] is applied to referential expressions only. Finally, both in the case of animacy and in the case of topicality, it is only the features of the possessor that are taken into consideration.

Various other relevant items were controlled as well, leading to further restrictions (Rosenbach 2002: 131–133). Thus possessors ending in /s/, /z/ or /θ/ are excluded, as well as all plural possessors. Finally, the style factor has been controlled: all texts come from stylistically comparable novels. It will be clear that, given all these restrictions and exclusions, Rosenbach’s aim to explain why the distribution of the ‘s-genitive and the of-genitive is the way it is (Rosenbach 2002: 1) can at best only be partly achieved. In particular the exclusion of the factor of complexity, and the restriction to one specific type of text only, severely explanatory potential of her study. Moreover, as in most statistical studies, no attempt is made to account for the exceptions to the tendencies observed.
4. The present study

4.1 A multifunctional approach

In the remainder of this paper, I would like to advocate a multifunctional approach to the question of what determines a speaker’s choice between a prenominal possessive and a postnominal of-construction which takes into account the various factors identified by traditional linguists. The present proposal will, however, go beyond these traditional accounts in assuming that

i. the various factors identified have the same cognitive-pragmatic origin;

ii. as a result, they typically supplement or reinforce each other;

iii. where two or more factors are in competition, the more basic factor(s) typically outweigh(s) the other factor(s);

iv. occasionally, one or more less powerful factors may overrule the more basic ones; in that case, the speaker’s preference can again be explained in terms of the same underlying cognitive-pragmatic principle.

As a multifunctional approach, my proposal contains elements of all three single-factor approaches described in the previous section. It will become clear that, however, that it has by far the most in common with that of Taylor (1996). Thus, in what follows, too, it will be assumed that the basic function of the possessive is to enable the hearer to identify the intended referent with as little cognitive effort as possible. Nevertheless, the present proposal differs from Taylor’s account in two important respects. First of all, according to Taylor, prenominal possessives differ from postnominal ones in that in the former the possessor is typically topical, or activated, while in the latter it is not. In addition, he argues, in most prenominal possessive constructions the possessee has not been explicitly mentioned in the text and will therefore have a higher degree of saliency than the possessor. As we will see, this characterization needs some qualification. Thus we find that in prenominal possessive constructions possessees need not be (textually) new, but are, in fact, often semi-active and sometimes fully activated and discourse-topical. Moreover, the possessor in such constructions need not be the discourse topic, and may even introduce a new entity into the discourse.

To account for these occurrences, I will assume that what is at issue here is not so much the topicality/activatedness of the possessor and possessee, but rather that of the relation between possessor and possessee. Generally speaking, what distinguishes prenominal possessives from postnominal of-phrases is that in the former the relation between the concepts denoted by possessor and possessee is presented as topical, or at least (semi-)active or inferable on the basis of some activated frame (general or specific knowledge structure), whereas in the latter construction this relation is created. This will account for the fact that in prenominal constructions the possessee is typically semi-active, as it will have been evoked through the activation of a knowledge structure in which the possessee is linked to the possessor. It will be clear that the more intrinsic or entrenched the relation between possessor and possessee, the more likely it
is for this relation to be activated through mention of the possessor, and, consequently, the more likely it will be for the possessor to appear in prenominal position. It is this very property which makes prenominal possessors ideal reference points: they are typically topical and identifiable, thus facilitating identifiability of the overall referent; but even if they are not, the hearer will be assumed to be aware of a relation between the possessor and some other (discourse-salient) entity; as a result, this other entity can be located within this knowledge structure, thus gaining a certain degree of identifiability.

A second difference with Taylor’s proposal is, of course, that, although the most basic of all, I do not regard topicality as the only principle determining word order in possessive constructions (cf. Mackenzie 1983:44ff.). Other factors, such as complexity, ambiguity and, occasionally, stylistic considerations, may also influence a speaker’s choice of construction. In most cases, these other principles will simply supplement or reinforce the basic principle and favour the same construction. In some cases, however, one or more of these principles may compete with the topicality principle; occasionally this may result in the topicality principle being overruled in favour of other principles.

4.2 Non-prototypical prenominal possessives: Some examples

In this section, we will look at a number of non-prototypical prenominal possessives, i.e. prenominal possessives which lack one or more of the features typically associated with this construction. Discussion of these examples enables us to examine how the various factors interact, and which factors (if any) are most dominant.

4.2.1 Indefinite possessors

Although typically definite, prenominal possessor can also be indefinite, as long as the indefinite NP can be assumed to function as a reference point for the hearer (see also Taylor 1996:187–193). Although use of an indefinite possessor does not lead to full (unique) identification of the referent of the construction as a whole, it may enable the hearer to locate the entity referred to in the vicinity of some associated concept, thus giving this entity a certain degree of identifiability. In most cases this is achieved by using an indefinite possessor which, although not referring to any identifiable entity, denotes a concept which may be assumed to be (semi-)active in the given context. This seems to be the case in the examples in (10).

(10)  a. A pass over a defender’s head calls for him to turn and play an awkward bouncing ball. <ICE-GB:W2D-015 #33:1>
  b. But whether you accept a witness’s evidence or not and what he or she says is entirely a matter for you <ICE-GB:S2A-061 #23:1:A>

In these examples, the possessor NP refers to a specific, but unidentifiable entity which clearly forms part of the mental representation the hearer has built up during the preceding discourse: (10a) is part of a coverage of a football match, (10b) is part of a judge’s summing-up in court. In both cases the concept denoted by the nouns defender and witness will be (semi-)active, as will the relationship between these concepts and
Evelien Keizer

the possessee nouns: in (10a), the relation is both intrinsic and activated by the given context; in (10b) the concept of evidence is evoked in the preceding discourse, as is its relation with the concept of witness.

In other cases, however, the indefinite possessor cannot be assumed to denote an active concept: in each of the examples in (11), the referent of the specific indefinite possessor NP is neither identifiable nor active/topical. In each case, however, the possessor is a person and the relation between possessor and possessee an intrinsic one, part of the frame, or general knowledge structure (GKS), activated through reference to a person irrespective of the context. As such, both the concepts denoted by the possessee nouns and their relation with the possessor can be assumed to be semi-active.

(11) a. POLICE hunting the killer of a bank manager’s “perfect” son were last night looking for a mystery jogger. <ICE-GB:W2C-020 #75:4>
   b. Sports fanatic David Nock, 16, collapsed dying into a policeman’s arms after a frenzied knife attack. <ICE-GB:W2C-020 #76:4>
   c. It was done with an unsure hand, but it was better than nothing, maybe better than she thought, because outside in the corridor a man’s head swivelled to watch her as his wife tugged on his arm. <ICE-GB:W2F-010 #87:1>

A similar explanation can be given of the use of a prenominal possessive in the examples in (12). These are again far removed from the prototypical prenominal construction, as the possessor is not identifiable, not specific, not human and, in the case of (12b), not activated. In (12a) the concept “apple” may be seen as semi-active, since the conversation mentions various fruits (lemons and pears have been referred to). Moreover, reference has been made to the texture of these fruits, thus preparing the way for a reference to “flesh” (which is, of course, also an intrinsic property of an apple). In (12b) neither the concept of “car” nor that of “starter motor” can be assumed to be active. Again, however, the relation is an intrinsic one (part-whole), thus justifying the use of a prenominal possessive.

(12) a. They’re not smooth like an apple’s flesh <ICE-GB:S1A-009 #166:1:A>
   b. With the Fairlite the Art of Noise were able to fulfil the futurist dream by making music from any sounds they could record including slamming doors, and a car’s starter motor <ICE-GB:S2B-023 #26:1:A>

In other words, indefinite (even non-specific, non-human) prenominal possessors, though not part of the prototype for prenominal possessive constructions, are acceptable if the relation between possessor and possessee can be assumed to be (semi-)active (or at least available). In that case, the indefinite possessor can serve as a reference point for the hearer, a function typically associated with – but not the prerogative of – definite, specific NPs. This means that it is not so much the definiteness (or specificity) of the possessor which determines whether or not it can appear in prenominal position, but rather its cognitive-pragmatic function in the given context.
4.2.2 First-mention proper names

Another interesting group of prenominal possessives is that in which the possessor takes the form of a first-mention proper name. As in such cases the proper name is unlikely to be topical or activated (although it may refer to a semi-active entity), Taylor’s claim that prenominal possessors are topical obviously does not apply here. However, if we assume that it is not the possessor itself that needs to be (semi-)active, but rather the relation between this possessor and the possessee noun, most of these cases become perfectly straightforward. If the proper name refers to a person, any concept intrinsically related to this person will be activated (to some degree). Where the relation is not an intrinsic one, it may still be activated on the basis of specific knowledge about this person, especially where this knowledge is relevant in the given context. In other words, apart from the fact that proper names are typically (uniquely) identifiable, proper names functioning as first-mention possessors are no different from non-topical indefinite possessors.

Consider the examples in (13), where the possessor is a first-mention proper name, introducing a new discourse entity. This is, in fact, a very frequent use of the construction, which is not surprising in view of the fact that it is a feature of proper names that they can be used to introduce an entity into the discourse that is familiar and (uniquely) identifiable for the hearer. In (13a), for instance, the name Bruce Hayes has not been mentioned before. The referent is, however, clearly assumed to be known to the hearer and so is the fact that he proposed an analysis. Since the relation between the possessor and the possessee is obviously relevant in the given context, it may be assumed to be a prominent feature of the knowledge frame activated by the use of the proper name, and may as such be assumed to be semi-active. Similarly, in (13b), Emma Clark is mentioned for the first time in the discourse; at the same time, however, the referent is clearly known to both participants, as well as semi-active in the context. Note that the concept denoted by the possessee, too, is textually activated.

(13)  a. and I think when you look at Bruce Hayes’s analysis, which John didn’t take very seriously <ICE-GB:S1A-005 #9:1:A>

b. A: By the way there’s a message for you. Call this person if you want this booking immediately.
   B: It’s my business. This is all this is Orla McGuinness … Have you got her number there so I can ring them now.
   A: Orla McGuinness, I’ve got Emma Clark’s number.
   <ICE-GB:S1A-074 #348:6:A>-<ICE-GB:S1A-074 #338:6:A>

4.2.3 Plural and/or inanimate possessors

The final category of non-prototypical uses of the prenominal possessive construction I would like to discuss is that of constructions containing a plural and/or inanimate possessor.

Let us begin with plural prenominal possessors. As pointed out before, there is, according to Huddleston and Pullum (2002:478), a small tendency in spoken language
to place plural NPs in postnominal position in order to avoid any potential ambiguity. As it turns out, however, cases of potential ambiguity are, in fact, quite rare. After all, if the prenominal possessor is typically topical, its number can usually be assumed to be known; in which case ambiguity is unlikely to arise and a plural NP is acceptable in prenominal position. In (14), for instance, the defendants and allies referred to in the prenominal possessor have been mentioned in the preceding discourse.

(14) a. From the defendants’ point of view however it doesn’t look that way at all

b. Our diplomatic editor Nick Gowing in London assesses Iran’s latest offer amid the allies' continuing bombardments

Inanimate possessors, too, often turn out to be quite acceptable in prenominal position. As argued before, animacy is not an independent factor in the choice for the prenominal possessive; instead, the preference for animate, human possessors in prenominal position derives from the fact that, more so than inanimate entities, humans tend to be topical and, as a result, activated in a given context.

However, since the basic requirement for the use of prenominal possessive is that the relation between possessor and possessee is activated, inanimate prenominal possessors are perfectly acceptable in situations where this requirement is fulfilled. In some cases, activation is due to the intrinsic nature between possessor and possessee (such as “part-of” or “property-of”, e.g. an apple’s flesh, a car’s starter motor in (12)); in other cases, the relation is context-specific, rendered (semi-)active through its relevance to the discourse topic.

Consider in this respect the example in (15). Here the possessee is clearly activated, as is the referent of the construction as a whole. It may even be argued that the possessor is not really needed here for the identification of the overall referent; instead, its presence seems to prepare the comparison with other corpora that is to follow. However, whatever the reason for mentioning the SEU, it will be clear that possessor, possessee and the relation between them can all be assumed to be activated.

(15) Describing the Corpus Parameter Tree

As I described above, the corpus parameters are related in a hierarchical structure, each parameter group splitting into further sub-groups in a tree-like manner. In the SEU’s parameter trees there are no loops between parameters, and it is equally unlikely that there would be loops in a similar parameter tree describing another corpus.

The strength of the topicality/activatedness factor is even clearer in those cases where a construction is non-prototypical in more than one respect. In (16), for instance, the possessors are both inanimate and plural. However, since the possessors are activated (through earlier mention or as part of an evoked frame) and the relation between the two concepts intrinsic (part-whole), this relation can be assumed to be (semi-)active.
Prenominal possessives in English

(16) a. The huge B Fifty-Twos arrived at Fairford last Tuesday. Eight of them have remained parked at the side of the runway all week. Maintenance crews have worked on them every day but this morning they began loading bombs and sweeping snow off the planes’ wings and fuselage.

b. And they break away. And Sibor very smartly into his stride, and he’s uh coming right over to these stands’ rails … Nayy Fasmal in the uh brown jacket but Sibor is right on the stands’ rails.

A similar account can be given of the prenominal use of possessors that are not merely inanimate, but which denote some abstract concept. In (17), for instance, both the proposal and the concept of support are clearly activated (through the mention of plan and support), as is the relation between them.

(17) I enclose a copy of a letter from the Chairman of the local CPRE about our problem with the Rural Area Local Plan. In para 3 John Smith recommends seeking the support of local councillors. More to the point is the identification of the proposal’s supporters.

One factor that is not easily outweighed by the topicality/activatedness factor, however, is that of complexity. Although this factor cannot by itself, on each occasion, account for the choice between a prenominal and a postnominal possessor, it seems to be the only factor which can overrule the topicality/activatedness factor. Where a possessor exhibits a limited degree of complexity, as in the examples in (18), the topicality/activatedness principle still wins out. Where the possessor has a high degree of complexity, containing, for instance, a phrasal or clausal postmodifier or an appositional NP, it must appear in postnominal position.

(18) a. This was Baghdad last night. Six hours of bombing in the Iraqi capital’s south-east suburbs.

b. I take it that the right honourable gentleman’s policy is to abolish the pound sterling, the greatest expression of sovereignty.

In many cases, however, the decision to place a possessor of some complexity in either prenominal or postnominal position is influenced by a number of factors. In those cases where the complexity principle competes with any of the other principles, the speaker will automatically be involved in some kind of balancing act, whereby such matters as the kind of text (e.g. spoken/written; formal/informal), the speaker’s assumptions concerning the processing capabilities of the addressee, as well as his/her own capabilities or stylistic preferences, may all influence the speaker’s choice.
4.3 Postnominal possessors: Some examples

Although the scope of this paper does not allow for a detailed discussion of postnominal possessives, it may be useful to look at just a few examples to illustrate the difference between prenominal and postnominal possessive constructions. However, due to the fact that discussion here is based on authentic material, it is not possible to compare alternative structures in the same context. It is, of course, possible, and in some cases certainly useful, to manipulate authentic data to show that an alternative construction would or would not have been acceptable in a given setting. The point is, however, that in the discourse context in question, the speaker chose to use one construction rather than another – in other words, the manipulated examples always lack an authentic discourse context and are as such less interesting for any linguist interested in the actual use of linguistic expressions. This does not mean that comparisons are not possible – it simply means that comparison will have to be made between different examples in their own discourse contexts.

It makes sense to assume that the use of postnominal constructions can be accounted in terms of the same set of factors that have been shown to play a role in a speaker’s choice for a prenominal possessive construction. Thus it is tempting to assume that the prototypical postnominal possessive is characterized by a maximally different cluster of factors from those identified for prenominal possessives; i.e. that postnominal *of*-constructions have a salient, possibly new, possessor, which may be definite or indefinite, specific or non-specific, singular or plural, human (animate or inanimate), with at least a certain structural complexity, while the relation between the two concepts is inactive and not intrinsic. As can be seen in (19), there are, indeed, constructions which exhibit all these characteristics:

(19) a. Professor Steve Humphries is the director of a British Heart Foundation research group based at the Sonly Centre in west London <ICE-GB:S2B-038 #29:1:A>

b. … The correct procedure is to write uh on that matter to the chairman of the Select Committee on Members’ Interests <ICE-GB:S1B-051 #46:1:C>

c. These of course are the famous opening words of Julius Caesar’s commentaries on the Gallic wars, uh, his, uh conquest of Gaul which took place during the fifties BC <ICE-GB:S2A-022 #4:1:A>

As pointed out by various linguists, it is equally true that there are very few cases in which the postnominal NP exhibits none of these properties, and, if such constructions do occur, they are generally considered marked or non-preferred (e.g. constructions like *the hair of John,* “the sister of me).

However, if we look at the way postnominal *of*-constructions are actually used, we find that in most cases one or more of these features is absent. Let us briefly consider two of such non-prototypical cases.
Prenominal possessives in English

(20) a. she's written seventeen books so far on Celtic mysteries uh Arthurian legends uhm, all kinds of stuff feminism the feminine aspect of God <ICE-GB:S1A-096 #99:1:A>

b. Jennings was fresh from drafting the constitution of Malaysia <ICE-GB:S2B-025 #74:1:A>

In (20a), the postnominal possessor, God, is a simple, singular, specific NP, referring to a unique being even higher up in the chain of being than humans. Nevertheless, use of a prenominal possessive would be infelicitous here. The reason is, once again, that the referent of the possessor NP has not been introduced, is not topical and cannot even be assumed to be semi-active in the given context. The function of the of-phrase is to enable identification of the possessee noun by linking this noun to an entity uniquely identifiable on the basis of long-term knowledge. Although the possessee noun is (semi-)active (through the mention of feminism), neither the possessor, nor the relation between possessor and possessee can be assumed to be activated – which is exactly what the use of a prenominal possessive would have suggested. In (20b), too, the possessor is not obviously related to the discourse topic. Moreover, although countries may be assumed to have constitutions, the construction in (20c) clearly has an existential quality: the constitution referred to does not, as yet, exist.

The possessors thus provide contextually new, focal information. It is on the basis of these pragmatic factors that the postnominal of-construction is preferred over the prenominal possessive.

Occasionally, even pronouns can appear in the non-prototypical, postnominal position. An example is given in (21):

(21) And after about three or four minutes of looking at the top half of her I pulled the covers back <ICE-GB:S1B-049 #47:1:B>

Here we find almost all the features which favour the use of a prenominal possessive: a human referent (the speaker’s new-born daughter) who is the topic of the conversation, a pronominal possessor and what looks like an intrinsic (body part) relation between possessor and possessee. Yet, a postnominal of-construction is not only acceptable here, but clearly preferred. The reason for this is that although the relation between possessee and possessor is one of part-whole, it is not really intrinsic: “top half” is not a body part that is normally evoked in relation to the human body. Nor has the concept of “top half” been introduced in the preceding discourse. As a result it cannot be assumed to have been activated and use of a prenominal possessive is out of place.

Unlike in the case of prenominal possessives, however, the choice for the postnominal construction does not seem to be determined by one major underlying principle (Keizer to appear); instead, each of the factors distinguished can be quite acceptably violated when outnumbered by competing factors. Consider, for instance, the example in (22).

(22) Good punch that from Mason his best yet in between the gloves of Lewis. Spears out his left jab again into the face of Mason once twice. Pokes out his
jab once two three times but only into the gloves of Mason who covers up well.

Here, the postnominal possessors take the form of definite, structurally simple NPs, with specific, human referents. Moreover, both referents are clearly activated in the given context (radio coverage of a boxing match), while the relation between possessors and possesses is intrinsic and active. However, although both Lewis and Mason are obviously discourse topics, sentence topicality shifts back and forth between them, according to the development of the fight and the point of view chosen by the commentator. Consequently, each occurrence of these names will be salient (although not contrastive). This means that in this case discourse salience overrules all other factors, resulting in a highly non-prototypical, but perfectly acceptable postnominal possessive construction.

5. Conclusion

In this paper I have used close text analysis of authentic examples in their context to demonstrate that a speaker’s decision to use a prenominal possessive (where a postnominal of-construction would have been semantically and syntactically acceptable) is determined, not by a single factor, but by a number of interacting and, to some extent, interdependent, factors. These, we have seen, include, at least, the following:

(23) i. topicality versus saliency of the possessor/possessee
    ii. activatedness of the relation between possessor and possessee
    iii. intrinsicness of the relation between possessor and possessee
    iv. “referent point” versus specifying function of the possessor
    v. complexity of the possessor
    vi. the gender/animacy of the possessor/possessee
    vii. number of the possessor
    viii. scope ambiguities
    ix. presence of certain types of pre- or postmodifier of the head noun
    x. stylistic considerations

Together, it has been argued, these factors allow us to define the prototypical prenominal possessive construction as consisting of a topical, active possessor and a salient possessee. The possessor is definite, specific and identifiable, functioning as a reference point for the identification of the possessee. In addition, the possessor has human

9. Note that contrastive stress on the possessor, whether in pre- or postnominal position, would suggests that one boxer is facing two opponents, who need to be distinguished from each other. Since, however, each boxer is facing only one opponent, confusion is unlikely to arise and contrastive stress is not needed.
reference, is singular and structurally simple. Finally, the relation between the two concepts is intrinsic and (semi-)active.

As will have become clear from the discussion in Section 4.2, there are, however, also many non-prototypical cases, where a prenominal possessor is used in the absence of one or more of these features. Close examination of these exceptions has resulted in two important generalizations. In the first place, it turns out that where several factors compete, the most powerful factors are complexity, and, in particular, the topicality/activatedness of the relation between possessor and possessee. Furthermore, it has been shown that these non-prototypical cases can be accounted for on the basis of the same cognitive-pragmatic principle that unites the factors mentioned in (23).

Unfortunately, it has not been possible to discuss in detail all the factors influencing a speaker’s choice for one of the two constructions. Nor has it been possible to offer a close examination of the use of postnominal possessives. Here, too, it has been suggested, a prototype can be defined, characterized by a cluster of factors maximally different from those identified for prenominal possessives. As pointed out in Section 4.3, however, postnominal possessives are not used in a very coherent fashion: unlike in the case of prenominal possessives, there does not seem to be one major underlying principle determining the choice for the postnominal construction; instead, each of the factors can be quite acceptably overruled by a combination of competing factors. Capturing the regularities and restrictions involved here therefore promises to be an interesting area for further research.

References


Ditransitive clauses in English with special reference to Lancashire dialect

Anna Siewierska and Willem Hollmann
Lancaster University

Unlike many other languages, English has three ditransitive constructions; a prepositional one in which the recipient or beneficiary is marked by a preposition and two double object constructions, one in which the recipient precedes the theme and the other in which the theme precedes the recipient. The availability and use of these three constructions in different dialects of English has long been an issue of controversy. This paper offers actual corpus data relating to the distribution of the three ditransitive constructions in one English dialect, namely Lancashire dialect. It shows that in cases where both the recipient and theme are pronominal, the double object construction with theme > recipient order is not only possible but in fact dominant.

1. Introduction

In this paper we address a complicated area of English grammar: the coding of the theme and recipient in ditransitive clauses. The literature, both descriptive and theoretical, reveals that there is quite some confusion with respect to the forms of encoding of the two constituents found in English. The confusion relates to the nature of the encoding patterns that are claimed to occur and also to the conditions underlying their occurrence. This confusion, we argue, can only be rectified by using a corpus-based approach. Anticipating a larger-scale study, here we take a small step in this direction, and consider the patterns of encoding of the objects in ditransitive clauses on the basis of corpus data from Lancashire dialect.

The paper is organized as follows. Section 2 presents what are considered to be the canonical forms of encoding of the non-subject arguments in ditransitive clauses in English and provides a brief overview of some of the major differences between the constructions in question. In Section 3 we consider the alternative realizations of the above constructions which have been noted in the literature, concentrating on the pronominal vs. nominal status of the non-subject arguments. Special attention will be given to the conflicting views of scholars with respect to which ditransitive patterns are in fact possible, which preferred and in which varieties of English. In Section 4 we will
Anna Siewierska and Willem Hollmann discuss the distribution and conditions of occurrence of the ditransitive patterns found in the dialect of Lancashire. The discussion will be based on corpus data extracted from four different corpora. And finally, in Section 5, we will seek to relate our findings with respect to the patterns of encoding in ditransitive clauses in Lancashire to some of the questions which have been raised in the theoretical literature pertaining to ditransitive clauses, as well as to more general issues in linguistic theory.

2. The two canonical patterns of encoding

English is one of the relatively few languages (see e.g. Siewierska 1998) in which the transfer of possession, either actual or intended, can be expressed by means of two truth-conditionally synonymous constructions, as illustrated in (1).

(1) a. John gave a book to Mary.
   b. John gave Mary a book.

The construction in (1a) is typically referred to as the dative or prepositional construction, the one in (1b) as the ditransitive or double object construction. In the prepositional construction the theme, book in (1a), occurs without prepositional marking and occupies immediate postverbal position, while the recipient, Mary in (1a), follows and is preceded by the preposition to – if the verb takes a benefactive rather than a recipient, as is the case with e.g. buy, fetch or find – for. In the double object construction the recipient is placed immediately after the verb with the theme following and neither evinces any prepositional marking.

Much ink has been spent on the syntactic, semantic and pragmatic characteristics of these two constructions. Syntactically the two are typically seen to differ with respect to grammatical relations, the recipient being an indirect object or under some analyses a direct object in (1b) but not in (1a) (see e.g. Ziv & Sheintuch 1979; Hudson 1992). Semantically, the double object construction is viewed as highlighting the transfer of possession, the prepositional construction the location of the transferred item (see e.g. Goldberg 1992). And pragmatically, the double object construction is associated with topical recipients and focal themes, the prepositional construction with topical themes and focal recipients (see e.g. Polinsky 1998).

Another factor differentiating the two constructions is their respective sensitivity to length and heaviness of the theme and recipient. English, like most other languages (see in particular Hawkins 1994), exhibits a preference for linearizing light constituents before heavy ones. This tendency is very much in evidence in the double object construction, which overwhelmingly features recipients consisting of a single word, typically a pronoun, but is much less strong in the prepositional construction. Biber et al. (1999) present data from the Longman Spoken and Written English Corpus.
Ditransitive clauses in English with special reference to Lancashire dialect

Table 1. Length of the recipient and theme in give, sell, offer (adapted from Biber et al. 1999:928)

<table>
<thead>
<tr>
<th></th>
<th>1 word</th>
<th>2 words</th>
<th>3+ words</th>
</tr>
</thead>
<tbody>
<tr>
<td>recipient-theme</td>
<td>recipient</td>
<td>85%</td>
<td>10%</td>
</tr>
<tr>
<td>theme</td>
<td>15%</td>
<td>35%</td>
<td>50%</td>
</tr>
<tr>
<td>theme-to recipient</td>
<td>to recipient</td>
<td>55%</td>
<td>25%</td>
</tr>
<tr>
<td>to theme</td>
<td>45%</td>
<td>30%</td>
<td>25%</td>
</tr>
</tbody>
</table>

(LSWE)\(^1\) supporting the above relating to the length of the theme and recipient with the verbs give, sell and offer in the two constructions; see Table 1.

We see that while 85% of the recipients in the double object construction are single words, the corresponding figure for themes in the prepositional construction is much lower, 55%. Nonetheless, a tendency for short-before-long linearization in the prepositional construction can also be discerned. Further evidence comes from the fact that themes which are heavy or complex may be postposed after the prepositional recipient. A case in point is (2):

(2) This irregularity in her features was not grotesque, but charming and gave to Anastasia’s face a humor she herself did not possess. (Biber et al. 1999:928)

It is important to note in this context that no comparable postposing is possible with heavy recipients in the double object construction. Examples such as (3) from Larson (1988:354) are invariably considered by syntacticians as ungrammatical.

(3) *John sent a letter every musician in the orchestra.

As the effects of weight on the encoding of arguments are so well known, it is obviously not weight that is the source of the confusion surrounding the encoding of the theme and recipient alluded to in the introduction. Accordingly, in what follows matters of weight will not be considered further.

3. Variation on the canonical patterns

Contrary to what is often supposed, the order of the theme and recipient, both in the prepositional construction and the double object one, can be switched. This seems to be least widespread when both of the constituents are full NPs, but consider:

(4) She |gave to her BRÓTHer a SÍGnet ‘ring| (Quirk et al. 1985:1396)

---

\(^1\) The LWSE is a 40 million word corpus of spoken and written English. The emphasis in the corpus lies on British English but American English is also represented. For more information see Biber et al. (1999:24–38).
The symbols representing tone unit boundaries (|) and rising/falling intonation (;) suggest that the pragmatics are important here. An instance of analogous switching in the double object construction is given in (5):

(5) She gave a book the man. (Hughes & Trudgill 1996:16)

Again, information structure probably plays a role, but also dialect: this variant “is not especially common, but does occur in northern varieties, particularly (…) if man is contrastively stressed” (Hughes & Trudgill 1996:16).

The permutability of theme and recipient appears to be somewhat more frequent when one or both of the non-subject arguments is/are pronominal. There is clear consensus (e.g. Quirk et al. 1985:1396; Larson 1988:364; Hughes & Trudgill 1996:16; Wales 1996:87; Cardinaletti 1999:61) that in cases with a pronominal theme and a full NP recipient the order recipient-theme is out, regardless of the presence or absence of to.

(6) *She gave (to) the man it.

This unacceptability is typically attributed to the clash between the topical character of the pronoun it and the focality associated with end position in English (see e.g. Quirk et al. 1985:1361 and passim; Polinsky 1998). Example (6) is also in breach of the short-before-long principle. However, if the pronominal theme is not a personal pronoun but a demonstrative or indefinite pronoun, as in (7), a full NP recipient is possible.

(7) a. They gave our guests that.
   b. I gave John some.

The ungrammaticality of the constructions in (6) leaves the prepositional construction in (8) and the double object configuration in (9):

(8) She gave it to the man.

(9) She gave it the man.

Freeborn (1995:205) notes a potential non-standard association with (9). Hughes and Trudgill suggest that the difference is primarily dialectal but also depends on speakers’ social (educational) background. They link the construction used in (8) with Standard English, and say that it is most common in the south of England, while “in the educated speech of people from the north [the pattern displayed by (9)] is also possible” (1996:16). Quirk et al. (1985) do not discuss any differences between these two possibilities – in fact they do not give any examples of either.

Moving on to cases with a pronominal recipient and a nominal theme, the double object pattern illustrated in (10) is more common than the prepositional construction in (11).

2. Unfortunately, this suggestion is not supported by references to other literature.
Ditransitive clauses in English with special reference to Lancashire dialect

(10) She gave him a signet ring. (Quirk et al. 1985:1396)
(11) She gave a signet ring to him.

The skewed distribution in favour of the recipient-theme order finds clear support in the LSWE: with give, offer and sell, the double object construction is four times as common as the prepositional pattern (Biber et al. 1999:928). And since 50% of the recipients in the relevant double object constructions in the corpus were pronominal as opposed to only 5% pronominal themes, it follows that the vast majority of the double object constructions involved pronominal recipients and nominal themes. Similar observations with respect to the preference for (10) over (11) have been made by Hawkins (1994:312) and Givón (1993, Vol. 2:219). It remains to be seen, however, whether it holds for all verbs. In any case, (11) is by no means rare.

We have not found any discussion in the literature of the pattern represented by (12), below, but informal enquiries among native British English speakers (from the North West) suggest that it is not entirely unacceptable, particularly if the theme carries contrastive stress:

(12) She gave to him a book.

About the fourth logical possibility, the double object construction with recipient and theme reversed, Hughes and Trudgill state that it “is not so common, but can be heard in [educated speakers in] the north of England, particularly if there is contrastive stress on him” (1996:16):

(13) She gave the book him. (Hughes & Trudgill 1996:16)

The highest degree of complexity – or perhaps we should say confusion – arises with twin pronominal objects. Of the four logical possibilities illustrated in (14)–(17), all but (15) are claimed to occur.

(14) She gave it to him.
(15) *She gave to him it.
(16) She gave him it.
(17) She gave it him.

Biber et al. (1999:929) suggest that the prepositional construction with theme-before-recipient (14) is the most common pattern in English, both inside and outside Britain. Quirk et al. (1985:1396) do not take a clear position. Kirk (1985), basing his observations on the Survey of English Dialects (SED), suggests that in Britain the prepositional construction is giving way to the double object one. Cheshire (1993:75) reports that the Survey of British Dialect Grammar (SBDG) reveals that the prepositional construction has been ousted by the double object one in many urban areas. According to Hughes and Trudgill (1996:16), in Standard English the double object recipient-theme

3. For more information about the SBDG see Edwards and Cheshire (1989).
order (16) is the norm. Koopman and Van der Wurff (2000:265) acknowledge that the prepositional pattern is common across varieties of English, though it is not seen as the dominant order – see further below. One is inclined to attach most credibility to Biber et al., as their suggestion is based on LWS E corpus data. Their statistics, overall clearly pointing to the prevalence of the prepositional construction over both of the double object ones, are shown in Table 2 below. The distribution of the constructions in question (all with pronominal recipients and themes) is represented as the number of instances per million words in four registers: conversation, fiction, news and academic prose.

In regard to the two double object patterns, the theme-before-recipient pattern is typically seen as a feature solely of British varieties of English. Apart from this, opinions are again divided. Quirk et al. (1985:1396) once more do not take a clear stance on the balance between the double object orders. Huddleston and Pullum (2002:248) suggest that theme-recipient is less common than recipient-theme. Biber et al. (1999:929) show that register is the deciding factor: recipient-before-theme is twice as frequent as theme-before-recipient in conversation, whereas in fiction the reverse pattern prevails (cf. our Table 2). To some degree in contrast with the findings of Biber et al., Koopman and Van der Wurff (2000:261), when discussing English in general, do not mention the recipient-theme order. It is commented on in their discussion of the dialectal distribution, where they suggest that while the south primarily has theme-recipient, “[s]ome northern varieties have the order IO-DO here, as do American and Australian English (though the ‘to’-phrase is probably more common)” (2000:265; emphasis original). Hughes and Trudgill seem to imply that in the South recipient-theme is widespread (though less common than the prepositional pattern cf. 1996:16), while the reverse (theme-recipient) pattern is “very common indeed” (ibid.) among educated northern speakers, but is “also quite acceptable to many southern speakers” (ibid.). Kirk (1985), on the other hand, associates the theme-before recipient order with the West Midlands and the South. Cheshire (1993:75) notes no clear regional preference for it, although she does identify a preference for the recipient-before-theme order in the urban areas of the South.

The literature survey above suggests that the dialects of the North of England exhibit a particularly rich array of encoding possibilities of the theme and recipient in ditransitive clauses. While both of the double object patterns are attested in the South and the North when the theme and recipient are pronouns, with mixed or two lexical constituents the North appears to have more variation in both the double object and

<table>
<thead>
<tr>
<th></th>
<th>Conv</th>
<th>Fiction</th>
<th>News</th>
<th>Academic</th>
</tr>
</thead>
<tbody>
<tr>
<td>theme-to recipient</td>
<td>90</td>
<td>70</td>
<td>10</td>
<td>&lt;5</td>
</tr>
<tr>
<td>recipient-theme</td>
<td>40</td>
<td>&lt;5</td>
<td>&lt;5</td>
<td>&lt;5</td>
</tr>
<tr>
<td>theme-recipient</td>
<td>20</td>
<td>10</td>
<td>&lt;5</td>
<td>&lt;5</td>
</tr>
</tbody>
</table>

Table 2. Distribution of ditransitive constructions with pronominal objects per million words (adapted from Biber et al. 1999:928)
Ditransitive clauses in English with special reference to Lancashire dialect

prepositional constructions. Furthermore, it is clear that the strongest contradictions in the literature obtain with respect to the distribution of the theme-before-recipient order in the double object construction in the North. Importantly, the questionnaire data from the SED or the SBDG are not delicate enough to allow these contradictions to be resolved. The data are somewhat dated by now, and they are also too restricted. The SED elicited only one ditransitive pattern for each participant: respondents were presented with the question *Jack wants to have Tommy’s ball and says to him, not: Keep it!* [fieldworker gesticulates]: . . ., cf. Orton (1962:100) and were only allowed to give a single answer. The SBDG did allow respondents to indicate that more than one response was acceptable (Cheshire et al. 1993:59) but again the only verb considered was *give*. In fact the whole discussion of the dialectal variation in the encoding of the theme and recipient in ditransitive clauses reported on above has been essentially confined to this verb. This is problematic: do the same patterns of encoding occur with other ditransitive verbs as with *give*? Are they confined to themes and recipients or do they extend also to themes and benefitaries? In the case of pronominal arguments, are the patterns in question displayed by all the possible combinations of pronominal themes and recipients/beneficiaries or only with a subset? Keeping these questions in mind, let us take a closer look at the ditransitives found in one of the areas of the North, namely Lancashire.

4. Patterns of argument encoding in Lancashire dialect

In our quest for Lancashire dialect data we relied on 4 corpora.

First, we used the British National Corpus (BNC). This is a 100 million word corpus of spoken and written Present-day English; for detailed information see e.g. Aston and Burnard (1998). The BNC contains 10 spoken texts categorized as Lancashire dialect, running to a total of about 150,000 words.

Second, we used the Lancashire texts of what we will refer to here as the Survey of English Dialects Incidental Recordings Corpus (SED-IRC). During the course of the SED project, in the 1950s and 1960s, many of the respondents were recorded by the fieldworkers. In total 289 out of the 313 SED localities are covered. Orton (1962:19) points out that the recordings in question were usually personal reminiscences or opinions. They vary between around 8 to 20 minutes in length. Both in terms of their structure and content as well as the range of respondents included, this aspect of the project seems to have been less than fully systematic. Reading the description in Orton (1962:18–19) one cannot escape the impression that these recordings were an

4. Siewierska (1998:179) raises the same question in a typological context: discussions of ditransitives in grammars of other languages tend to be restricted to *give* as well.

5. We are grateful to Andrew Hardie and Tony McEnery for clarifying several issues related to the recordings in question, and their electronic transcription.
afterthought. At any rate, the original 78 rpm records and reel-to-reel tapes are kept at the University of Leeds. Under the supervision of Juhani Klemola (Leeds) the recordings were transcribed. The transcribed data have not been made publicly available yet. Lancaster University was involved in the (part-of-speech) tagging process. The Lancashire texts we considered are from the following localities: Coniston, Cartmel, Dolphinholme, Eccleston, Fleetwood, Harwood, Marshside, Read and Ribchester. The absence of some localities that are included in the printed SED should be seen in the light of the shortcomings in the design of the corpus in terms of rigorousness, described above (for the full list of Lancashire localities see Orton 1962:31). The total size of SED-IRC is approximately 800,000 words; the Lancashire part is around 22,000 words.

Third, we searched the Freiburg English Dialect Corpus. The compilation of FRED is led by Bernd Kortmann. The project is still in progress. When complete, this spoken corpus will amount to almost 2.5 million words, from dialects from all over England (including Wales) and Scotland. Transcription of the Lancashire files is complete; they amount to around 250,000 words (23 texts). For more information about the corpus the reader is referred to the FRED home page: http://www.anglistik.unifreiburg.de/institut/lskortmann/FRED/ [31 May 2004].

Fourth, we retrieved the relevant data from the Lancashire part of the Helsinki Corpus of British English Dialects (HD). Having been initiated by Harold Orton and Tauno Mustanoja in the 1970s this project is also still ongoing, currently under the supervision of Kirsti Peitsara. The complete corpus will contain more than 1 million words of spoken data from various parts of England, mainly East-Anglia, the South West, Essex and Lancashire. The Lancashire material (from Barrowford, Colne and Nelson), which was collected by Riita Kerman in the 1980s, amounts to a little more than 60,000 words if the fieldworker’s speech is also counted, and a little under 50,000 words if it is not. More information about the corpus can be found on the www page: http://www.eng.helsinki.fi/varieng/team3/1_3_4_2_hd.htm [31 May 2004].

From the above corpora we isolated all the occurring instances of ditransitives by searching for all English verbs listed by Levin (1993:45–47) as displaying dative and/or benefactive alternations and then manually identifying the instances of actual ditransitive use.

In all we found 449 ditransitive clauses featuring the following verbs: *blow, bring, buy, cook, draw, fetch, get, give, hire, make, offer, owe, pay, pay, post, read, save, sell, send,*

6. We should like to express our gratitude to Bernd Kortmann for granting us access to the Lancashire data.

7. We would like to thank Kirsti Peitsara for running the searches for ditransitives on the corpus for us.

8. This figure includes only clauses containing themes and recipients/benefactives which could potentially occur both in the prepositional and double object constructions. So for example it includes *He told this to me* which alternates with *He told me this* but not *He told me that he wasn’t going to come back*; cf. *He told that he wasn’t going to come back (to) me.*
show, take, teach, tell and write. The vast majority of the ditransitive clauses, 83 per cent (N = 374), were double object ones. This is very much in line with Biber et al’s (1999) findings as to the clear dominance of the double object construction over the prepositional one with respect to the verbs give, send and offer (it was mentioned earlier that they detected a skewing of about 80 per cent as against 20 per cent). This dominance of the double object construction obtained for all noun/pronoun combinations found in the data (but see further below). Among the 449 examples of ditransitive clauses there were 34 clauses where both of the relevant constituents were nominal, 338 clauses with mixed nominal and pronominal constituents and 77 with two pronouns. In the first group, the double object construction occurred in 73 per cent of the clauses (N = 25), in the second group in 86 per cent (N = 292) and in the third group in 74 per cent of cases (N = 57).

Turning to the heart of the matter, the placement of the theme and recipient/benefactive relative to each other in the respective constructions, all the instances of the prepositional construction were of the canonical type, i.e. with theme-before-recipient/benefactive order. In just under half of the cases (N = 37) the theme was a pronoun and the recipient/benefactive a lexical NP. There were, however, 9 instances such as those in (18) of the converse pattern.

(18) a. They brought the check for you. (Freiburg)
    b. Send this letter to them. (BNC)
    c. You can buy a ticket for them for seven pounds. (Helsinki)

The canonical pattern was also overwhelmingly dominant among the double object clauses; 94 per cent (N = 353) had recipient/benefactive-before-theme order as opposed to only 6 per cent (N = 21) with theme-before-recipient/benefactive order. What is, of course, crucial with respect to the two ordering possibilities in the double object construction is the categorial breakdown of the theme and recipient/benefactive. With two nominal objects, all instances involved the placement of the recipient/benefactive before the theme. With a mixed combination of one nominal and one pronominal constituent recipient/benefactive-before theme order occurred in 286 instances and theme-before-recipient order in 6. Thus the canonical pattern is again dominant accounting for 98% of the cases. As might be expected, in virtually all instances of the recipient-first order it was the recipient that was pronominal. The three exceptions are shown in (19).

(19) a. Show your father them. (Freiburg)
    b. Give Alex one. (BNC)
    c. Show Sid one. (BNC)

Note that (19a) represents an instance of (6), a pattern considered in the literature to be categorically excluded. The six instances of theme-before recipient order all involved pronominal themes as in (20).
(20)  
  a.  I’ll give it your sister. (BNC)
  b.  So he gave it Tom. (SED)
  c.  Give it all the kids either side. (Freiburg)

So far our data have not lent much substance to the claims concerning the strong presence of theme-before recipient order in ditransitives in the dialects of the North of England. Recall from the survey of the literature in Section 3, however, that the claims with respect to the theme-before recipient order are strongest in relation to two pronominal arguments. In order to see whether they do indeed hold, we must be particularly careful in the nature of the constituents that we consider. Among the 57 instances of the double object construction where both the theme and recipient/benefactive are pronouns, there are 34 in which the pronominal theme is either non-specific some, one, any, anything, or the demonstrative this or that. In all the instances in question the theme follows the recipient/benefactive. If we include these in our considerations, then again the canonical recipient-first order emerges as dominant over the theme-first order, the relevant figures being 73 per cent (N = 42) vs. 27 per cent (N = 15). It is preferable, however, to define pronominal in a more narrow sense, to include personal pronouns only. The reason is that especially in the historical literature on English it has often been observed that it is specifically personal pronoun objects that may display special syntactic behaviour as compared to nominal objects (Smith 1893; Bacquet 1962; Mitchell 1964:119, 1985:§§3889, 3907; all cited by Koopman 1994:9). Recall also the difference in the acceptability of pattern (6) depending on whether the pronoun is or is not a personal one. If we restrict our attention, then, to clauses containing two personal pronouns, the total number unfortunately falls below 30 (i.e. N = 23), meaning that statistical significance is compromised. To the extent that we are still allowed to draw some tentative conclusions, it is interesting to note that we now find the converse situation; the theme-first order prevails over the recipient-first order, 65 per cent (N = 15) vs. 25 per cent (N = 8). Thus when both theme and recipient are personal pronouns, it appears that the placement of the theme before the recipient is indeed the preferred pattern in the double object construction in the investigated dialectal data from Lancashire. In fact, the theme-before recipient order is not only dominant in the double object construction but dominant overall, as it also obtains in the prepositional construction. In all, in 81% of the clauses with both a pronominal theme and recipient/benefactive the former precedes the latter. It is important to note in this connection that the narrowing down of our focus to include only personal pronouns entails the disappearance of the obvious preference for the double object construction over the prepositional one with two pronominal participants; the relevant figures are 58 vs. 42 per cent as opposed to the previous 74 vs. 26 per cent.

Our data is presented in Tables 3–7, below. Tables 3–6 tease the data apart for each of the 4 corpora; Table 7 conflates the results from all of them. The figures in parentheses in Tables 3–7 refer to the number of all pronouns, i.e. not only personal ones, but also unspecified forms such as someone or anyone, the proform one or demonstratives.
Table 3. Distribution of the complementation patterns in ditransitives in the Lancashire part of the BNC

<table>
<thead>
<tr>
<th></th>
<th>2 Pro</th>
<th>ProNP</th>
<th>NPPro</th>
<th>NP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>RT</td>
<td>4 (21)</td>
<td>71</td>
<td>(2)</td>
<td>7</td>
<td>82 (101)</td>
</tr>
<tr>
<td>PP</td>
<td>11 (13)</td>
<td>16</td>
<td>4</td>
<td></td>
<td>31 (33)</td>
</tr>
<tr>
<td>Total</td>
<td>21 (40)</td>
<td>89</td>
<td>4 (6)</td>
<td>7</td>
<td>121 (142)</td>
</tr>
</tbody>
</table>

Table 4. Distribution of the complementation patterns in ditransitives in FRED

<table>
<thead>
<tr>
<th></th>
<th>2 Pro</th>
<th>ProNP</th>
<th>NPPro</th>
<th>NP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>RT</td>
<td>4 (19)</td>
<td>171</td>
<td>1</td>
<td>12</td>
<td>188 (203)</td>
</tr>
<tr>
<td>PP</td>
<td>5</td>
<td>16</td>
<td>1</td>
<td>6</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>15 (30)</td>
<td>189</td>
<td>2</td>
<td>18</td>
<td>224 (239)</td>
</tr>
</tbody>
</table>

Table 5. Distribution of the complementation patterns in ditransitives in the Lancashire part of SED-IRC

<table>
<thead>
<tr>
<th></th>
<th>2 Pro</th>
<th>ProNP</th>
<th>NPPro</th>
<th>NP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>RT</td>
<td>(1)</td>
<td>3</td>
<td></td>
<td></td>
<td>4 (11)</td>
</tr>
<tr>
<td>PP</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>3 (4)</td>
<td>14</td>
<td></td>
<td>17</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 6. Distribution of the complementation patterns in ditransitives in the Lancashire part of HD

<table>
<thead>
<tr>
<th></th>
<th>2 Pro</th>
<th>ProNP</th>
<th>NPPro</th>
<th>NP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR</td>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>RT</td>
<td>(1)</td>
<td>31</td>
<td>6</td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>PP</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>34</td>
<td>4</td>
<td>6</td>
<td></td>
<td>44</td>
</tr>
</tbody>
</table>

Table 7. Distribution of the complementation patterns in ditransitives: Combined results

<table>
<thead>
<tr>
<th></th>
<th>2 Pro</th>
<th>ProNP</th>
<th>NPPro</th>
<th>NP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR</td>
<td>15</td>
<td>6</td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>RT</td>
<td>8 (42)</td>
<td>283</td>
<td>3</td>
<td>25</td>
<td>319 (353)</td>
</tr>
<tr>
<td>PP</td>
<td>20</td>
<td>37</td>
<td>9</td>
<td>9</td>
<td>75</td>
</tr>
<tr>
<td>Total</td>
<td>43 (77)</td>
<td>326</td>
<td>12</td>
<td>34</td>
<td>415 (449)</td>
</tr>
</tbody>
</table>
The above data reveal that there appear to be three as opposed to two possibilities of the encoding of the non-subject arguments in ditransitives only when the theme is a personal pronoun and the recipient a noun or when both are personal pronouns. In the first instance the prepositional construction is favoured over either of the double object patterns (37 vs. 9); among these 9 double object patterns the theme-first pattern is favoured over the recipient-first one (6 vs. 3). In the second instance, as mentioned earlier, there seems to be a very slight preference for the double object pattern over the prepositional one (23 vs. 20), but again a favouring of the theme-first pattern in the double object construction over the recipient-first one. Given the above, the question arises what determines the choice of encoding. Two obvious factors to consider are the nature of the verb and the person/number/gender features of the pronouns.

Of the verbs involved in both dative and benefactive alternations that we considered only three occurred in the corpus in the alternative double object pattern, i.e. the theme-before recipient one. The verbs in question are give, send and show. To the best of our knowledge this is the first time the existence of the alternative double object pattern has been attested with verbs other than give. That the verbs in question are all dative-alternating rather than also benefactive-alternating may or may not be significant. In any case, we rather suspect that there are more verbs displaying this variation. Give, send and show are among the most frequently occurring ditransitive verbs. It is therefore not altogether surprising that in a small corpus such as ours, the fairly uncommon theme-before-recipient pattern fails to show up with less frequent verbs. However, we would expect a larger corpus to yield a greater variety of verbs displaying the theme-first double object pattern. Needless to say, all three verbs, give, send and show, also occurred in the other two patterns. Therefore in the case of at least these three verbs, it is not the nature of the verb that underlies the choice of ditransitive pattern. As for the features of the pronominal theme, since first and second person pronouns feature only rarely as themes in ditransitive clauses, it is not surprising that in

9. In fact various languages are claimed to manifest a prohibition against first and second person themes in double object clauses if co-occurring with third person pronominal recipients. This is especially common in the case of pronominal clitics as, for example, in Italian (Cardinaletti 1999:64) or Polish or in the case of weak forms as in Swedish. The relevant constraint is illustrated below from Polish in which a third person dative may be followed by a first person accusative only if the latter is a full pronoun. Compare (i) with (ii).

(i) Pokazala mu/jej/im mnie
    showed:3SGF him/her/them:DAT IACC
    ‘She showed me to him/her/them.' ?She showed him/her/them me.’

(ii) *Pokazala mu/jej/im mi
    showed:3SGF him/her/them:DAT IACC
    ‘She showed me to him/her/them.' ?She showed him/her/them me.’

Note also the oddity of the double object as opposed to prepositional construction in such cases in English.
all three ditransitive patterns the theme was virtually always third person and nonhuman. The five instances of themes that were human, two me, one us, one him and one them, all occurred in the prepositional construction. All the other pronominal themes were thus either it or them, with it predominating (46 vs. 19). There is little to add about the properties of the non-pronominal recipients in the relevant patterns. Recall that in the three instances of the canonical double object construction with a nominal recipient and a pronominal theme, the recipient was a single word. Somewhat surprisingly, if one disregards the preposition, there were relatively more two-word and longer recipients in the alternative double object construction than in the prepositional one. However, the number of instances of the alternative double object construction is too small to allow for meaningful generalizations to be made.

Turning to the patterns with two personal pronouns, we note that recipients, unlike themes, tend not to exhibit person restrictions. Accordingly, the corpus exhibits the full range of human personal pronouns functioning as recipients, first, second and third (masculine and feminine) both in the singular and plural, as shown in the examples in (21).

(21) a. Give me it to go away.
    b. He wants to sell us it.
    c. I’ll give you them.
    d. So send it me.
    e. I’ll give it you.
    f. Give them us.
    g. Give them to me.
    h. We sent them to you.
    i. I taught it to them.

Nonetheless there are evident differences in the distribution of the person/number combinations of the recipients in the three ditransitive patterns found in the corpus. The canonical ditransitive pattern exhibits a dispreference for third person recipients, there being only one instance, namely the him/it combination shown in (22a).

(22) a. When they showed him it.
    b. I gave it him.
    c. Give it her for next time.

In both the alternative double object pattern and the prepositional one, there is no such dispreference. In the alternative double object construction we find three instances of it/him as in (22b) and two of it/her as in (22c) though none of it/them, which in turn is attested four times in the prepositional construction. The alternative double object construction evinces the highest percentage (33%) of the it/me combination. The above differences may be just coincidental, due to the small number of ditransitives with two personal pronouns of each type. Nonetheless the possibility cannot be excluded that the two double object patterns disfavour certain combinations of pronouns, because of case recoverability problems or due to phonological factors. Case
recoverability problems are most likely to arise when both of the pronouns are animate, a situation not attested in our corpus. With respect to phonological factors, it is often claimed (see e.g. Larson 1988:364) that in the canonical double object construction a pronominal recipient preceding a pronominal theme must be necessarily unstressed. If this is indeed so, then both *him* and *them* may emerge as ‘m, leading to ambiguity particularly when followed by *it*. Assuming that in the alternative double object construction the necessarily unstressed pronominal is again the immediately postverbal one, i.e. the theme rather than the recipient, the distinction between *him* and *them* in an *it/him* or *it/them* combination should be maintained. This could account for the apparent paucity of third person combinations in the canonical double object pattern as opposed to the alternative one. Interestingly enough, in neither of the double object patterns attested in the corpus do we find a *them/them* combination.

In sum, our analysis of the order of the theme and recipient in ditransitive clauses has shown that in the investigated dialect of Lancashire, when both theme and recipient are personal pronouns, there is a clear preference for positioning the theme before the recipient. This order obtained in 81% (35/43) of the relevant instances. The preference for theme-first order is met not only by the use of the prepositional construction, as in the standard language, but also by the use of the alternative double object construction. Of the 35 instances of theme first order, 43% (15) are by means of the alternative double object construction. Significantly, when both the theme and recipient are personal pronouns, the alternative double object construction is nearly twice as common as the canonical one, the relevant figures being 65% vs. 35%. Thus while the most common construction with two personal pronouns is the prepositional one, the alternative double object construction and not the canonical double object one is the next in line.

5. Theoretical relevance of findings

The existence of three as opposed to two ditransitive patterns in some dialects of English, even if only with respect to certain types of themes and recipients/benefactives, raises a number of interesting questions. First of all, do the theme and recipient/benefactive in the two double object patterns manifest the same or different grammatical relations? In other words, does the change in order reflect a mere change in say pragmatic status or a more fundamental structural difference? Given the restricted nature of the theme in the alternative double object construction, it is rather difficult to answer the above, as the standard syntactic tests used for distinguishing indirect from direct objects in English, for instance asymmetries in binding, extraction and right-
ward movement (however interpreted), cannot be applied. Another set of questions, also bearing in part on the issue of grammatical relations, concerns the morpho-phonological form of the pronouns in the three constructions. According to the typology of structural deficiency recently developed by Cardinaletti and Starke (1999), pronominal forms may be divided into strong, weak and clitic forms. Only strong forms may be coordinated and modified. Weak forms or clitics cannot. Weak forms, unlike clitics, however, may bear word stress and may be deleted under ellipsis. They do not, on the other hand, form clusters. What then is the status of the pronominal themes and recipients in terms of this typology? Both the theme and recipient/benefactive in the prepositional construction can be modified and coordinated as well as separated by a verbal particle, as shown in (23), and thus emerge as strong forms.

(23) a. He gave them all to me and her.
   b. He gave it back to me.

But it is by no means obvious what the status of the relevant forms in the canonical double object construction, let alone the alternative one, is. Is it indeed the case that the inner pronoun in both constructions can never be stressed and thus qualifies as a clitic rather than a weak form? What are the modification or coordination possibilities of the outer pronouns in the two constructions? The outer theme can be modified (He gave us it all) and coordinated (He gave me both her and him) but whether the same holds for the recipient in the alternative double object one we simply do not yet know. The morpho-phonological status of the pronominal theme and recipient in the two double object patterns is of considerable interest as this may have a direct bearing on how the two patterns are to be dealt with in a model of grammar. In various syntactic frameworks alternative ordering patterns involving clitics or weak forms are handled by quite different mechanisms than those involving strong forms or lexical NPs. Accordingly, there are likely to be fewer consequences for the grammar of ditransitives if neither of the two pronominals in the two double object patterns emerge as strong than if a strong form is involved in any of the two patterns.

Another, more general, issue, which has been implicit in the discussion above, concerns the study of variation in language itself, more specifically the importance of taking variation seriously (Croft 2001). English ditransitives have been treated in much of the literature as a more or less clear-cut class. As our summary in Sections 2 and 3 of the findings of Biber et al. (1999) indicates, in their section on ditransitives they do not distinguish between American English and British English. A variation on this lumping approach is displayed by Goldberg (1992). In her analysis of the grammatical function of the theme in the double object construction she follows Dryer (1986) and Bresnan and Moshi (1990) in arguing that it is a different type of object than a direct object. In a footnote, she suggests that the evidence for object status is partly constituted by direct object semantics, the absence of a preposition, but also the passivizability of the theme (as well as of the recipient) in British English (Goldberg 1992:71, n.4), cf. It was given him. Moving on to the evidence against it being a direct
object, she observes that the theme always follows the recipient when both are present and that passivizability does not obtain in American English (1992:71, n.4).

Goldberg's claim about the relative order of recipient and theme cannot be upheld in the light of the data presented above, which show that there is variation. In addition, she bases her unitary grammatical analysis (of the theme) on an aspect of syntactic variation across American and British English. Croft (2001) argues that in view of cross-linguistic variation in the instantiation of what are traditionally seen as universal syntactic categories and constructions it is a mistake, methodologically speaking, to assume the existence of these categories. Instead, syntactic structure is language-specific. It does not follow from this that there is nothing universal in languages: the constraints on the mapping between form (phonological, morphological and syntactic structure) and function (semantic and discourse-pragmatic meaning) are universal (Croft 2001:e.g. 61). Therefore, linguists should direct their search for language universals to the form-function mapping.

To get back to the case of ditransitives, the apparent variation between American and British English leads us to conclude that a distinction must be made between the American English double object construction and the British English one. Both constructions should be investigated in their own right. To the extent that Goldberg's generalization concerning the passivization facts is valid, they may indicate that the status of the theme in the British English recipient/benefactive-theme double object construction is different from that in the American English one. And concerning the constructions as a whole, the question arises whether perhaps in British English it is somehow more highly transitive than in American English (the connection between passivizability and semantic transitivity is widely accepted, see e.g. Bolinger 1978; Keenan 1985; Rice 1987; cf. Siewierska 1984 for a critical evaluation of this position). Another, possibly additional, explanation might be that the British English passive construction is wider in scope than the American English one.

Having made – in line with Croft (2001) – the case for distinguishing between American English and British in the study of ditransitives (and the passive, and indeed constructions in general), we would in fact go even further. The Lancashire data suggest that even a language-specific double object construction is too simplistic. The form-function mapping in ditransitives in regional dialects should not necessarily be expected to conform to that of the standard variety, and indeed it does not, as is shown most clearly by the theme-recipient variant of the double object construction. Now while regional dialects may be socially or politically stigmatized, from the linguist’s point of view there is nothing that should make them any less valid or interesting as sources of information about the universal constraints on the mapping between form

11. In addition, Croft (2001) takes the position that semantic structure (speakers’ conceptual knowledge) is also pretty much universal. This is one of the basic assumptions of the semantic map approach to grammatical knowledge (Croft 2001: e.g. 92ff., 2003:133–139; Haspelmath 2003).
and function.\textsuperscript{12} Thus, for example, the function of, and restrictions on, the use of the theme-before-recipient/benefactive order are not only relevant in connection with our understanding of the role of discourse pragmatics in argument linking and/or the distinction between strong, weak and clitic pronouns in Lancashire dialect. Instead, they will have a bearing on our understanding of these issues in language in general. In keeping with the spirit of Kortmann (2003), then, we argue that functional-typological linguistics should not abstract away from, but instead embrace the wealth of variation displayed across dialects. Such a position is very much in line with the desiderata for a truly functional grammar presented in Butler (2003) and espoused in Functional Discourse Grammar as outlined in Mackenzie (2003). This keen attitude towards dialectal variation will characterize our follow-up research on ditransitives. In our more in-depth analysis of the issues raised above (as well as additional questions that may arise) we intend to rely on several sources of information. We have already made a start on retrieving data from corpora. In addition to the ones used so far – outlined in Section 4, above – we are in the process of compiling a corpus specifically of Lancashire dialect. This is based on recordings held by the North West Sound Archive.\textsuperscript{13} The purpose of the Archive is to “record, collect and preserve sound recordings relevant to life in the North West of England” ([http://www.gmcre.co.uk/other/NWSA/nwsa.htm](http://www.gmcre.co.uk/other/NWSA/nwsa.htm) [NWSA www page, 28 May 2004]). It has more than 100,000 recordings of speech, many of them personal reminiscences. From these recordings we have initially selected around 60 hours of speech for transcription, the localities of the speakers matching, as much as possible, those of the informants of the Survey of English Dialects. This data will be important in shedding more light on issues such as the relative frequencies of the various alternative argument linking patterns in Lancashire. Our new corpus data may also shed more light on the range of verbs displaying the comparatively rare pronominal theme-before-recipient order, although in this connection we will be using written questionnaires and oral tasks as well. The same applies to the question as to whether different patterns co-exist in the grammar of a single speaker. As for the status of the pronouns (and thus, indirectly, the question of the validity of Cardinaletti and Starke’s (1999) structural deficiency typology involving the distinction between strong vs. weak vs. clitic pronouns) the value of corpus data has been, and is expected to remain, relatively limited. The double object patterns are not frequent to begin with, and are considerably less likely to occur with the structures indicative of their status, viz. coor-

\textsuperscript{12} The ultimate logical conclusion of this line of reasoning goes yet further: we should not expect individual speakers’ grammatical knowledge to be exactly the same even if they have the same dialect. The suggestion is that a speaker’s knowledge is a function of the sum total of their unique linguistic experience (cf. also Bybee 2000; Croft 2000:26), although, again, the constraints on the form-function mapping in idiolects is universal.

\textsuperscript{13} We are very grateful to Andrew Schofield for helping us in selecting and copying our materials, and also to Andrew Hamer for helping us become familiar with the Archive.
dination and modification. It would not be feasible to construct a corpus sufficiently large that these patterns would show up in suitably high numbers. Passivization data, which may merit a closer look in connection with Goldberg's claim cited above, are likewise too rare for corpus data alone to be enough. In this connection then, our methodology will follow a recent trend in dialect grammar research towards triangulation (e.g. Cornips & Jongenburger 2001; Cornips & Poletto 2005). Cornips and Poletto (2005:941–945) point out that a purely observational (corpus-based) approach runs the risk of remaining blind to certain aspects of language use, especially uncommon patterns, while a solely experimental method (written questionnaires and oral tasks) also has shortcomings, e.g. the well-known tendency for subjects who speak a non-standard variety to be influenced by prescriptive norms in various ways (cf. also e.g. Labov 1972:21, 177, 213, 1996:78). With roots in many branches of linguistics – ranging from typology to dialectology, and corpus linguistics to syntactic theory, including cognitively-oriented approaches – the results should be relevant across much of the spectrum of the field, although perhaps especially the broadly functional side, given the centrality of variation.

References

Ditransitive clauses in English with special reference to Lancashire dialect


'It was you that told me that, wasn’t it?’

*It*-clefs revisited in discourse*

María de los Ángeles Gómez-González
University of Santiago de Compostela

This corpus-based investigation explores the features and discourse distribution of *it*-clefs. It is argued that their core meaning is the positive identification of a discourse element, usually a subject/agent NP, mostly by means of the declarative pattern *it is/it was … that* in spontaneous persuasive speech. The construction is found to implement three main strategies: (a) corrective when reformulating old topics, generally displaying the pattern new + given; (b) transitional when (re)introducing new or deactivated topics or spatio-temporal settings, usually with the pattern new + new; and (c) topical when continuing with a previous discourse topic, normally exhibiting the values given + new. Lastly, the central function of *it*-clefs is shown to be that specifying a relation of exhaustive topicality.

1. Introduction

Based on evidence provided by the *International Corpus of English-Great Britain* (ICE-GB), this paper revisits the topic of *it*-clefs (ICs) (e.g. *It was you that told me that,*  

* This paper belongs to a bigger project entitled *Discourse Analysis in English: Aspects of cognition, typology and L2 acquisition* (http://ietsil.usc.es/scimitar), sponsored by the Spanish Ministry of Science and Technology (‘Ministerio de Ciencia y Tecnología’, MCYT), the European Regional Development Fund (‘Fondo Europeo de Desarrollo Regional’, FEDER), and the Xunta de Galicia (XUGA) (grant numbers BFF2002-02441, PGIDIT03PXIC20403PN). The support of these institutions is hereby acknowledged. I am also grateful to Chris Butler, Lachlan Mackenzie, Knud Lambrecht, Peter Collins, Francis Cornish, and Francisco González for very helpful comments and discussion on earlier drafts of this paper. A slightly modified version of this work has appeared in the *Southwest Journal of Linguistics*, 23(2), 81–120.
wasn’t it?). Traditionally, clefts1 – such as those in (1) below – have been said to differ from corresponding more basic, frequent or canonical structures – those in (2) – formally, semantically and functionally, but to be otherwise equivalent in terms of truth-conditions and illocutionary force (Borkin 1984; Davidse 2000; Dik 1978, 1997; Halliday 1967a, b, 2004; Quirk et al. 1985; f).

(1) a. It was you that told me that, wasn’t it? (CE-GB:S1A-099 #272:2:A)
   b. it is in the least developed countries that they have suffered the most <, >
      (ICE-GB:W1A-014 #30:1)

(2) a. You told me that, didn’t you? (ICE-GB:S1B-024 #93:1:C)
   b. (i) In the least developed countries they have suffered the most.
      (ii) They have suffered the most in the least developed countries.

Broadly speaking, clefting identifies a discourse strategy whereby information is packaged or “cleft” into two units in order to fulfill a two-fold discourse effect: (i) to set up a relationship of identification of the specifying type between two units; and (ii) to give discourse prominence to (part of) one of the two units.

My analysis will be restricted to it-clefts (ICs), leaving aside such constructions as wh-clefts, or pseudoclefts, and other cognate sequences, exemplified in (3) and (4) below, which differ in information structure, function and discourse effects (cf. Collins 1991; Delin & Oberlander 1995; Hedberg 1990; Prince 1978; Weiner & Miller 1996):

(3) a. What makes them go into politics is a character defect.
   b. Where they have suffered the most in the least developed countries.

(4) a. The (only) thing/reason that makes them go into politics is a character defect.
   b. The (only) place in which/where they have suffered the most in the least developed countries.

1. The term cleft (construction) was introduced by Jespersen (1965 [1927]:251) to designate those structures with the peculiarity of being ‘split’ or divided into two separate parts with the purpose of singling out one of these parts, which resulting directs attention to it and marks some sort of contrast. The label it-cleft, in turn, became widespread in the 70’s to refer to the subtype under analysis here (Hankamer 1974; Prince 1978). In Systemic Functional Grammar, on the other hand, they are called Predicated Theme constructions (Halliday 2004), while in Functional Grammar they are considered as a subtype of Focus constructions and are referred to as Dummy clefts (Dik 1997:292).

2. The examples extracted from the corpus are identified with the following abbreviations: (i) ICE-GB (the name of the corpus); (ii) S (‘Spoken’) or W (‘Written’); (iii) the text code (S1B), the text code number (024), the text unit number and a second number indicating the subtext (#93:1), and in spoken the speaker’s identification (e.g. #93:1:C).
Specifically, we shall explore the three types of constraints that characterize ICs, i.e.
formal (Section 3), semantic (Section 4) and functional (Section 5), in interaction
with such competing motivations as:

i. focus of attention (FA), that is, the camera angle of discourse that shifts from
one attentional window to another as discourse progresses, imposing different
directions of mental scanning (or different paths of mental access) in the
conceptualizers’ minds as language is construed and thereby evidencing the speakers’
subjectivity and point of view (cf. e.g. Cornish 2004; Langacker 2001a, b);

ii. the semantico-pragmatic notion of focus and the assumptions it leads to in terms
of presupposition and assertion (cf. Bolkestein 1998; Dik 1978, 1997; Drubig 2003;
1996);

iii. the mappings of given and new information, as well as the cohesive ties entailed by
such mappings (Connolly 1998; Dik 1978, 1994; Halliday & Hasan 1976; Gundel
1977, 1999, 2002; Martin 1992);

iv. topic, that is, a prominent conceptualization which acts as a kind of cognitive an-
choring point with respect to which other conceptualizations are brought into the
discourse (Kemmer 1995:58), at roughly two possible levels, a global discourse
level (global Topic) and a local discourse level (local Topic) (Bolkestein 1998; Dik
1978, 1997; Downing 1991, 1997);

v. the staging of information, referring to how speakers/writers initiate discourse
(units) and build them up incrementally as discourse unfolds, thereby directing
the conceptual monitoring of the message at issue (cf. Gómez-González 2001,
2004; Hannay 1994; Mackenzie 2000, 2004); and

vi. text-type, characterized by a specific communicative function and such contextual
variables as mode (roughly the medium through which linguistic contact occurs,
e.g. written, spoken), tenor (kind of speaker-addressee relationship, e.g. formal,
informal) and field (subcategories of style referring to the social processes in which
language plays a part) (Halliday 2004; Martin 1992).

Only a cursory description of the categories in (i–vi) above is offered here because
a more fine-grained account of the problems involved in their definition and ap-
lication, as well as the inter-analyst disagreements regarding these issues, has been
given elsewhere (Gómez-González 2001, 2004; see also e.g. Cornish 2004; Hannay
& Bolkestein 1998). In what follows we shall see how the formal, semantic and func-
tional characteristics of ICs in combination with the factors listed above explain their
distribution across the ICE-GB text-types (Section 6).

It will be argued that the core meaning of ICs is the positive, rather than the neg-
ative or contrastive identification of a discourse element, the element in-focus (EIF),
usually a subject/agent NP, mostly preceded by the declarative pattern it is/it was . . .
that, in mainly informal or spontaneous spoken language. We shall also offer an al-
ternative formal analysis of the construction, in particular with regard to the status of
it, which is interpreted as a contextual indexical item with both forward-looking and
backward-looking potential, as well as of the dependent clause, which is taken to be subordinated to the entire thematic predicative pattern it is/it was EIF (rather than rankshifted within the EIF) (cf. Huddleston 1984:460–462, 1988:185–186) and to display more freedom of relativization than what has often been admitted in the relevant literature (cf. Downing & Locke 1992:249; Quirk et al. 1985:1386–1387).

Furthermore, we shall see that ICs implement three main discourse strategies: (i) corrective (when reformulating or correcting old/previous topics, generally displaying the pattern new + given); (ii) transitional (when (re)introducing new or deactivated topics or spatio-temporal settings, usually with the pattern new + new); and (iii) topical (when continuing with a previous discourse topic and normally exhibiting the values given + new). Underlying these three strategies, it will be concluded, is the central function of IC, namely to direct, redirect or correct the addressee’s attention towards the EIF, which the speaker judges to be the current topic of discourse. For this reason it will be finally claimed that ICs convey an implication of exhaustive topicality, rather than one of exhaustiveness, as often postulated in previous work (cf. Ball 1994; Biber et al. 1999; Borkin 1984; Collins 1991; Davidse 2000; Declerck 1983, 1984, 1988, 1994; Delin & Oberlander 1995; Dik 1978, 1997; Geluykens 1988; Gundel 1977; Hedberg 2000; Huddleston 1984; Lambrecht 1994, 2001; Quirk et al. 1985; Ward, Birner, & Huddleston 2002; Weinert & Miller 1996).

2. Material and methods

As already stated, this study is based on the British component of the International Corpus of English (ICE-GB) (for further details see http://www.ucl.ac.uk/english-usage/ice). ICE-GB displays a major mode classification into spoken language (300 texts, approximately 600,000 words) and written language (200 texts, approximately 400,000 words), comprising a wide variety of texts. ICE-GB is released with the ICE Corpus Utility Program (ICECUP3), a text analysis programme especially devised to work with the tags and parse trees included in the corpus.

ICE-CUP3 selected 430 tokens of grammatical constructions tagged as CLEFTIT by the ICE-GB compilers. However, after a close scrutiny of these tokens, some of them had to be discarded from the analysis, for two main reasons. Either they turned out to be not ICs (e.g. Was it not the case (...) that he was (...) in the process of expansion, ICE-GB:S1B-064 #140:1:A, vs. that he was (...) in the process of expansion was not the case, which is not an IC, but a Subject extraposition); or they were incomprehensible or incomplete sequences resulting from contextual or processing disruptions (e.g. it was the style of humour of the day which, ICE-GB:S1-022 #75:1:C; it’s autosuggestion that if I might use that word, ICE-GB:S1B-070 #95:1:B; it’s really foreign taste is the, ICE-GB:S2A-057 #49:1:A).

Once the exclusions were made, our sample came up to a total of 422 ICs. Their frequencies and features will be discussed in the sections that follow. As a last methodological remark, note that frequencies result from dividing the number of examples by
the number of words in ICE-GB, 600,000 words in the spoken part and 400,000 in the written part, and then multiplying the quotient by 10,000 words.

3. Formal features

Our findings show that unmarked declarative ICs (383 cases, 90.76%) are more frequent than marked declaratives (5 tokens, representing 1.18% and occurring in a frequency of 0.08 per 10,000 words) or interrogatives (34 examples amounting to 8.06%, respectively). Figure 1 below displays the variations of the unmarked declarative pattern of ICs consisting of (i) a Head clause (A) containing a copular or relational verb (usually be) (It's a character defect), and (ii) a dependent clause (B) (that makes them go into politics).

Figure 1 provides a graphical representation of three important assumptions made here about the IC formal configuration: (i) 1 and 4 are presented as elements of a discontinuous description constituted by it + RDC; (ii) the RDC is characterized as an immediate constituent of the IC as a whole, dependent on the be-clause; and (iii) the wh-form of the RDC refers back to the unit dubbed EIF. Explanations for these three issues will provided in what follows as we present our findings for each of the four prototypical elements of the construction.

The subject of the superordinate clause (A: 1), it, has been interpreted in the literature roughly in two different ways: (i) either as a non-referential dummy element that acts as syntactic filler of the subject slot, or (ii), the position endorsed here, as a referential element of some kind. There seem to be two main reasons adduced by the supporters of the dummy-pronoun interpretation. One is the lack of agreement between it and the verb in the RDC. The assumption is that since the latter agrees with the EIF rather than with it (e.g. and it is such organisations that are more likely to be successful in periods of rapid change, ICE-GB:W2A-011 #40:1), then it cannot be a referential pronoun but must be a dummy one (cf. Collins 1991; Huddleston 1988). Declerck and Seki (1990) come to the same conclusion on the grounds that neither it can always be formulated as a nominal – which in their view means that there is no antecedent for the pronoun – nor can it be replaced by s/he or other pronouns when referring to human antecedents (cf. also Sornicola 1988:358–359; Weinert & Miller 1996 for similar conclusions).

In this study, however, we shall adopt the view that it is “referential” in the sense that it points both backwards and forwards in discourse. As a result of its backward-looking potential the it + be combination provides a link between the identifying construction it instantiates (Section 4) and what has come before. In Section 5 we shall see that this link is mainly used to mark topic (dis)continuity in discourse and/or to express the speaker’s point of view against the background of a context. In addition, exploiting its forward looking potential, it acts as a cataphoric pronoun that points to the variable represented by X or the “gap” in the open proposition encoded in the dependent clause (B) (e.g. “the X that makes them go into politics”, in the example
Note: The asterisk (*) symbolizes the optional tense and polarity markers that may also be housed within (X: 2), such as markers of aspect and modality, as well as focusing adverbs (only, just, etc.), as in (7a) and (8b), with the constraint that only one occurrence of each kind is possible (cf. Biber et al. 1999:959; Ward, Birner, & Huddleston 2002:1416).

Figure 1. Unmarked declarative patterns of ICs in ICE-GB
in Figure 1 above), the value of which is identified with the element EIF (a character defect) through the predicative construction it + be (further details on the semantic configuration of it-clefts will be given in Section 4).

This analysis has the advantage of explaining why it cannot be replaced by other pronouns, one of the arguments adduced by dummy-it supporters: the third person singular neuter pronoun is best suited to pointing forward to a variable (X), whose value is unknown (it may turn out to be animate or inanimate, feminine or masculine, singular or plural, etc.), in an open proposition, which crucially is not a referential expression proper. Concomitant accounts supporting some degree of referentiality in the it of it-clefts are provided by for instance Hedberg (2000:906, 819) and Gundel (2002:118), who assume that the cleft pronoun (it) together with the cleft clause (our RDC) functions as a discontinuous definite description, or Miller and Weinert (1998), who refer to it as a "definite deictic" (cf. also e.g. Akmajian 1970; Borkin 1984; Davidse 2000).

Although absent from ICE-GB, (A: 1) realizations other than it have also been attested in the literature, as illustrated in (5) below:

(5) a. (No,) that was the doctor I was speaking to (Quirk et al. 1985:1386)
   b. Those are my feet you’re treading on (Quirk et al. 1985:1386)

It would seem, however, that these instances differ from canonical ICs in some respects (cf. e.g. Ball 1977; Lambrechts 2001). First, the sequences in (5) are used specificationally – as opposed to predicational constructions (e.g. That's a nice dress you're wearing) – but unlike other specificational clefts those in (5a–b) cannot be uncleft, thereby resembling proverbial clefts which often cannot be uncleft (e.g. It’s the rare linguist who can keep track of these distinctions vs. ?The rare linguist can keep track of these distinctions) (cf. e.g. Ball 1978; Declerck 1988; Hedberg 1990, 2000). And second, the (A: 1) elements in (5), that, those, seem to be more referential than the it of it-clefts. For one thing, only the former can be used exophorically, thereby participating in an equation between two truly referential denotata. Thus, in the case of (5a), for example, we could utter the sequence when pointing to someone in front of us meaning that ‘that was the doctor, the person I was speaking to’. By contrast, the it of ICs cannot be used exophorically and neither can it participate in an equation between two truly referential denotata because the dependent clause (A) – as already noted and as will be shown in more detail in Section 4 – does not denote an entity or referent, but an open proposition.

Now, moving to (A: 2), the copular verb is generally be (99.29% of cases in ICE-GB), although it may also have other/periphrastic realizations such as look (like) or appear to (be), as exemplified in (6) below (0.71% in ICE-GB):

(6) a. It uh looks like Justin Channing who’s receiving treatment <,> (ICE-GB:S2A-003 #60:1:A)
b. Thereafter it appears only to be the very grandest of Sacramentaries that would use uncial for portions of the text other than headings and titles, such as the Drogo Sacramentary (ICE-GB:W2A-008 #56:1)

Table 1 below shows that (A: 2) has a preference for affirmative polarity, 92.9% with a relative frequency of 3.9 per 10,000 words, and for simple tenses, more precisely, either the simple present (65.87% with a frequency of 2.78) or the simple past (32.70% with a frequency of 1.38). Negative polarity (7.10% with a frequency of 0.3) and other tenses (1.41% with a frequency of 0.06) are only marginal. These variations are exemplified in (7) and (8) below.

(7) a. Simple Present: Uhm however what we say is that it is never one individual agency or one person that is responsible if a child is abused and left at home or taken into care and hasn’t been abused (ICE-GB:S1B-030 #110:1:E)

b. Simple Past: You just guessed that it was sort of a gradient that did that (ICE-GB:S1A-008 #85:1:A)

c. Future: It’ll be Neil Webb who goes across to take it with Brian McClair waiting at the near post (ICE-GB:S2A-003 #23:1:A)

d. Modal (Simple Past) Perfect Infinitive: It occurred to me that it might have been the contrast with Isabel that Barnsley had found attractive (ICE-GB:W2F-011 #92:1)

e. (Semantic) Future construction (with the Progressive be going to): As they come into the final bend now it’s going to be Chris Louis who takes the chequered flag (ICE-GB:S2A-012 #84:5:A)

(8) a. So, it’s not the people that are covered (ICE-GB:S1B-058 #94:1:E)

b. It was often not the governing class, but agrarian reformers or popular leaders who carried out annexations (ICE-GB:W2A-001 #34:1)

c. So , my theory is with these very different demographic patterns that you see within one pastoral population , shows that it cannot be pastoralism per se that is determining the demographic parameters here , but that you’ve got something else (ICE-GB:S2A-047 #132:1:A)

A closer consideration of negative ICs in ICE-GB reveals that most of them (27 out of 30 tokens) place the negative marker in the main clause (e.g. (8a) above), whereas there are only three cases in which negation is placed in the RDC (e.g. It was me who didn't want the physical relationship with Jeremy, ICE-GB:S1A-050 #89:1:B). This tendency seems to suggest that, if present, negative polarity in ICs tends to be focused upon with the most frequent purpose of “correcting” a discourse topic, as shall be further detailed in Section 5.2.2.

In addition, it was observed that negative ICs are more frequent in spoken texts (20 tokens (66.66%)/600,000 words, with a frequency of 0.33 per 10,000 words) than in written texts (10 tokens (33.33%)/400,000 words, with a frequency of 0.25 per 10,000 words). Despite their limited nature, these results could be said to support Tottie’s
Table 1. Distribution of tense and polarity in the COPULA of clefts in ICE-GB

<table>
<thead>
<tr>
<th>ICE-GB</th>
<th>TOTAL</th>
<th>Frequencies per 10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive Polarity</td>
<td>Negative Polarity</td>
</tr>
<tr>
<td>Simple Present</td>
<td>264 [62.55%]</td>
<td>14 [52.17%]</td>
</tr>
<tr>
<td>Simple Past</td>
<td>122 [28.90%]</td>
<td>16 [17.39%]</td>
</tr>
<tr>
<td>Future</td>
<td>3 [0.71%]</td>
<td>3 [0.71%]</td>
</tr>
<tr>
<td>Conditional</td>
<td>2 [0.47%]</td>
<td>2 [0.47%]</td>
</tr>
<tr>
<td>Pres. Progressive</td>
<td>1 [0.24%]</td>
<td>1 [0.23%]</td>
</tr>
<tr>
<td>TOTAL</td>
<td>392 [92.89%]</td>
<td>30 [7.10%]</td>
</tr>
</tbody>
</table>

Frequency (per 10,000 words)

"It was you that told me that wasn’t it?"
(1982: 88–105) claim that negation is more common in spoken language than in writing. We shall come back to the issue of negation in ICs when exploring the semantic and functional features of the construction in Sections 4 and 5, respectively.

Be that as it may, the formal simplicity exhibited by the copula of ICs reflects its comparatively secondary functional load: to set up a background of affirmative or less frequently negative identification against which a discourse element is brought to the fore. I dub this the element in-focus (EIF), that is, the element that becomes the focus of attention (FA) at a certain point in discourse as a result of the cleaving strategy. Its nature, as seen in my data, confirms the results of other investigations, that is, the preponderance of NPs (including proper nouns and pronouns), followed by PPs as the most frequent EIFs (e.g. Biber et al. 1999; Borkin 1984: 136; Collins 1991; Givón 1990: 704; Ward, Birner, & Huddleston 2002: 1418).

As shown in Figure 1 above, as well as in Table 2 and the examples in (9) below, in ICE-GB the EIF offers a range of five possible realizations, which may be either simple or complex (as in e.g. it was originally the driving worm and the wheel system (...) which was missing, ICE-GB: W2A-040 #104:1):

(9) a. NP: It’s the writer that gets you so involved (ICE-GB: S1A-016 #238:1:D)
    b. PP: It was only in December last year that the European Council decided to establish a new bank (ICE-GB: S1B-054 #19:1:D)
    c. AdvP: It is here that the democrats have made significant headway (ICE-GB: S2B-006 #18:1:C)
    d. FCl: It is when temples then became bigger and you stop using it that you can actually reduce the angle (ICE-GB: S2A-024 #86:1:A)
    e. NFCl: If that policy is then adopted by a Chief Constable it is he adopting that policy that is accountable for that policy that he adopts in his force area (ICE-GB: S1B-033 #80:1:C)
    f. Zero: It was you that told me that, wasn’t it? (ICE-GB: S1A-099 #272:2:A)

In both spoken and written texts the most common EIF is by far a noun phrase (NP) (76.77%) as in (9a). Remarkably, sixty tokens (1.84%) out of 324 are unmodified proper nouns (e.g. So it’s Hollywood Harry that goes on with Mack the Knife on the right, ICE-GB: S2A-006 #183:5:A), a possibility barred from restrictive relatives, as shall be explained below; and these are more common in speech than in writing: 54 tokens appear in spoken texts with a frequency of 0.9 per 10,000 words, and only 6 in written texts with a frequency of 0.15. This result seems to suggest that the use proper noun EIFs is associated with informal or non-standard language.

Second in frequency in both the spoken and the written mode come prepositional phrases (PPs) (15.16%). The preposition that appears in a larger number of cases is in (17 times), as in (9b), followed by until (8 times), with (7 times), and from (5 times), while other prepositions are more infrequent.

Adverbal phrases (AdvPs) are less common EIFs (4.73%), especially in written texts (with a frequency of 0.12 vs. 0.25 in the written mode). The adverb that appears in a larger number of cases, both in written and in spoken texts, is here (7 examples),
<table>
<thead>
<tr>
<th>S</th>
<th>DO</th>
<th>PrepC</th>
<th>Total</th>
<th>Frequencies per 10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NP</strong></td>
<td>271</td>
<td>6</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>227 (79.09%) / 600,000 words 3.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>64 (15.16%) / 1,000,000 words 0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36 (56.25%) / 600,000 words 0.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28 (43.75%) / 400,000 words 0.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PP</strong></td>
<td>0</td>
<td>61</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>36 (56.25%) / 600,000 words 0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15 (23.08%) / 400,000 words 0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 (1.56%) / 600,000 words 0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AdvP</strong></td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>15 (56.25%) / 600,000 words 0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 (19.61%) / 400,000 words 0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fin Cl</strong></td>
<td>1</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>5 (50%) / 400,000 words 0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NFin Cl</strong></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1 (100%) / 600,000 words 0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zero</strong></td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>273</td>
<td>96</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>188 (68.86%), 3.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>57 (20.28%), 0.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26 (9.75%), 0.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 (3.58%), 0.16</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
as in (9c) above, followed by then (5 times) and why (3 times), as opposed to other adverbs that appear just once (e.g. there, only, simply, etc.).

Clausal EIFs (Cls) are fourth in terms of frequency (2.59%). They may be either finite clauses (FinCl) (2.36%), mostly adverbial clauses of time, as in (9c), purpose, reason, means, and condition (with only if), or non-finite (NFinCl) (0.23%). The scarcity of clausal EIFs in both spoken and written texts (with a frequency of 0.10 and 0.12, respectively) is due to processing constraints whereby long or wordy units tend to be placed last to facilitate encoding and decoding processes. Interestingly, two tokens of clausal EIFs correspond to the so-called inferential sentences (i.e. It is not only that our criminal laws are being broken more profligately than ever, ICE-GB:S2A-039 #55-58:1:A; Is it just that the food is becoming boring <>, ICE-GB:S1A-059 #105:1:A). These are ICs that have a sentential EIF but lack the dependent clause (B). For an extensive discussion of this type the reader is addressed to e.g. Delahunty (2001) and references therein.

Lastly, zero realizations (0.71%) tally with three instances of interrogative IC, which lack both the EIF and the RDC because they are directly retrievable from the context, as in the tag question in (9f) above. Adjectival EIFs (e.g. It isn’t obsessive that Bill is, just manic-depressive, Delahunty 1984:76) have not been attested in ICE-GB, which presumably reinforces their dialectally marked condition (cf. Collins 1991:64).

Now considering syntactic functions and as a side-effect of the formal tendencies just described, EIFs equivalent to subjects (S) in uncleft constructions are most frequent (64.69%), followed by adjuncts (A) (22.74%), direct objects (DO) (7.34%) and prepositional complements (PrepC) (4.50%). These tendencies are reproduced in both written and spoken texts, although DO EIFs seem to be comparatively less frequent in written texts (only 5 cases (16.12%)/400,000 with a frequency of 0.12 per 10,000 words) than in spoken language (26 tokens (86.87%)/600,000 with a frequency of 0.43 per 10,000 words). EIFs playing the role of indirect object (IO), subject complement (SC) and object complement (OC) are also possible but these did not appear in ICE-GB. Examples of the seven categories are displayed in (10):

(10) a. S: It’s probably only half the population that are covered (ICE-GB:S1B-058 #56:1)
b. A: It was in one of the office buildings that I discovered the letters (ICE-GB:S2B-023 #61:3:A)
c. DO: It was actually the study of architecture I really enjoyed (ICE-GB:S1A-034 #16:1:D)
d. PrepC: It’s a sort of abdication of the responsibility of shaving that I’m most interested in, rather than the aesthetic sort of quality of the thing (ICE-GB:S1A-097 #285:1:A)
e. SC: It is angry that he was (Collins 1991:64)
f. OC: It’s dark green that we’ve painted the kitchen (Quirk et al. 1985:1385)
g. IO: It was to Terence (that) she made her appeal (Collins 1991:62)
Let us now focus on the dependent clause (B). This may be realized as a non-finite clause (NFCl), either gerund-participial or infinitival (1.65% in ICE-GB, e.g. It’s Lord Chalmer coming clear by six inside the final furlong, ICE-GB:S2A-006 #91:2:A) (cf. Ward, Birner, & Huddleston 2002:1420); but this realization is merely a deviation from the main pattern in which B is realized by a relative-like dependent clause (RDC) (98.35%).

As already advanced, B is represented in Figure 1 above as an immediate constituent of the main clause, dependent on the be-clause, rather than rankshifted or embedded within the EIF. The implication is that the referent of the EIF reveals the identity of the relative pronoun of the RDC as a dominant coreferential entity (in a hypotactic construct), rather than as being restricted in meaning by the relative post-modifier (for a similar account see Huddleston 1984:460–462, 1988:185–186). This analysis explains why the verb of the RDC generally agrees with the EIF – when the wh-form is the subject, although there may be some exceptions to this general tendency (e.g. It wasn’t him it was me who was at fault <>, ICE-GB:S1A-050 #84:1:B) – rather than with it: the whole unit dubbed EIF is the antecedent of the wh-form in the RDC, as also emphasized by e.g. Aarts (1997:212) and Davidse (2000:1107–1114). The former claims that ICs provide a sort of constituency test because only those strings of words that behave as units can appear in the EIF slot (e.g. It was a beautiful song that Louis sang yesterday vs. *It was a beautiful that Louis sang yesterday). Davidse, in turn, remarks that while in restrictive relatives the antecedent is the nominal head without the determiner, in RDCs the antecedent is the whole EIF unit.

Moreover, resembling the non-restrictive type of relatives (and reverberating their intonational patterns), the RDC elaborates its antecedent as information that is separate from, and secondary to, the remainder of the dominant clause. And it is due to this “secondary” status that the RDC may be omissible in some contexts, as has already been noted and will be confirmed by our corpus data below. However, RDCs are closer to restrictive relatives in that they can be introduced by a relative phrase (that, who, where, how and when in ICE-GB) or by zero, but differ from them mainly because RDCs allow much more freedom in relativizing (e.g. PPs, unmodified proper nouns) and in omissibility of the relative pronoun. For a summary of the main differences between cleft constructions and relative clauses, the reader is referred to e.g. Ball (1994), Collins (1991:52ff.), Gómez-González (2001:305–306), Huddleston (1984:460ff.) and Sornicola (1988:346), and the references therein.

In closing this section, let us analyse the formal realizations and distribution of the relative pronouns of ICs in ICE-GB, included in Table 3.

Table 3 shows that both in speech and writing the most frequent realizations of the wh-form of the RDCs are in decreasing order: that (56.16%), who (22.03%), which (12.79%) and zero (7.10%). That is by far the most common, as also noted by e.g. Quirk et al. (1985:1386–1387), which demonstrates the all-purpose nature of this relative pronoun in it-clefts. That is used after antecedents of all kinds: (i) NPs, as in (9a); (ii) pronouns (e.g. it was he that left, ICE-GB:S1A-075 #26:1:B); (iii) proper nouns (e.g. So it’s Hollywood Harry that goes on with Mack the Knife on the right, ICE-GB:S2A-006
Table 3. Formal realizations and distribution of the relative pronouns of it-clefts in ICE-GB

<table>
<thead>
<tr>
<th></th>
<th>that</th>
<th>who</th>
<th>which</th>
<th>Zero</th>
<th>Ø RDC</th>
<th>when</th>
<th>where</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>156 (65.82%)</td>
<td>71 (76.34%)</td>
<td>27 (50%)</td>
<td>26 (86.66%)</td>
<td>5 (100%)</td>
<td>1 (50%)</td>
<td>1 (100%)</td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>1.18</td>
<td>0.45</td>
<td>0.43</td>
<td>0.083</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Written</td>
<td>81 (35.18%)</td>
<td>22 (23.66%)</td>
<td>27 (50%)</td>
<td>4 (13.34%)</td>
<td>0</td>
<td>1 (50%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2.025</td>
<td>0.55</td>
<td>0.67%</td>
<td>0.1</td>
<td></td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>237 (56.16%)</td>
<td>93 (22.03%)</td>
<td>54 (12.79%)</td>
<td>30 (7.10%)</td>
<td>5 (1.18%)</td>
<td>2 (0.47%)</td>
<td>1 (0.23%)</td>
</tr>
</tbody>
</table>

With regard to which, interestingly enough, our findings question the claims made by e.g. Downing and Locke (1992:249) and Quirk et al. (1985:1386–1387): the former note that it is not possible to use which as the relative element of RDCs, while the latter find the same constraint when which, whose or whom are preceded by a preposition, in which case – they conclude – we would have a postmodifying relative clause. There are 54 tokens of which-RDC in our corpus, some of which are reproduced in (11) below:

(11) a. It was work which she much enjoyed (..) (ICE-GB:S2A-062 #8:1:A);
    b. It’s mainly the content which differs rather than the actual language itself (ICE-GB:S1B-003 #66:1:B)
    c. (..) it was this matter on which I consulted with the chairman of the Select Committee (ICE-GB:S1B-054 #33:1:B)

In four cases which is preceded by a preposition, as in (11c) above. The latter are indeed ambiguous sequences because, depending on the intonation pattern, they may be interpreted as it-clefts (e.g. (..) it was this MATTER on which I consulted with the chairman of the Select Committee) or as relative clauses (e.g. (..) it was this matter on which I CONSULTED with the chairman of the Select Committee). Nevertheless, ambiguities notwithstanding, our results demonstrate that ICs with which-headed RDCs do exist in English.

Zero realizations, in turn, are not so frequent as reported in other investigations, according to which that and zero are the prototypical relative elements in RDCs, instead of other wh-words (e.g. Quirk et al. 1985). In ICE-GB zero relatives (7.10%) are fourth in terms of frequency, after that, who and which – as already pointed out – and they are more frequent in speech (86.34% with a frequency of 0.43 per 10,000 words) than in writing (13.34%, with a frequency of 0.1). Moreover, only in spoken language can we find instances of zero relatives when the relativized element is functioning as subject (e.g. And it’s the grey Altaya still has it from Young Pretender in the yellow (..),
ICE-GB:S2A-006 #176:4:A), which indicates the informal or non-standard use of such realizations.

A subtype of zero realization is that in which both the relative element and the verb form of the RDC are elided because they are retrievable from the context (e.g. 'It's only going to be the Donaldson part of the library open probably' (ICE-GB:S1A-069 #213:2:A); But it's Great Britain still in possession (ICE-GB:S2A-004 #379:1:A); while an extreme version of this type corresponds to those ICs that have no dependent clause. Only 5 tokens of these have been recorded in ICE-GB: they can either be inferential ICs (when having a clausal EIF) – as already noted above – or reduced ICs (cf. Declerck & Seki 1990:15–51). In reduced ICs the RDC is again recoverable from the context: in ICE-GB all are the second conjoin of a coordinative construct sharing the RDC with the first conjoin (e.g. (...) So is it you Marjorie that’s got the results of our questionnaire or was it Mike [that’s got the results of our questionnaire], ICE-GB:SIB-077 #81:1:A).

Lastly, the occurrences of when (2 cases) (e.g. (...) it was only a few years ago when I said it to someone (...), ICE-GB:W1B-003 #142:2) and where (1 case) (i.e. it's this kind of routine work where she says her concentration is most affected, ICE-GB:S2B-011 #73:1:E) are only marginal. These results show that when there is a choice between one of these two pronouns and that, the latter is preferred over the other two thereby confirming its multi-purpose nature (e.g. (9b) or (9c) above).

Before closing this section, we must characterize the two deviant patterns attested in ICE-GB: marked and interrogative ICs. Interestingly, as illustrated in (12) below, in all marked ICs a proper noun is fronted to pre-subject initial position, that is, before it, and all but one (i.e. (5a–d)) are used by the same speaker (A) at different stages of the same spoken text-type (S2), a live sports commentary:

(12) a. Gary Stevens it was who come forward (ICE-GB:S2A-001 #97:1:A);
    b. Allan Smith it is who's got it (ICE-GB:S2A-001 #105:1:A);
    c. Simeone it is who chips the ball (ICE-GB:S2A-010 #174:1:A);
    d. Leram it is who's holding on to the lead (ICE-GB:S2A-012 #143:7:A);
    e. General Lawrence of Arabia it was who said that the press (...) (ICE-GB:SIB-031 #20:LA)

The examples in (12a–d) above show that by fronting the EIF the speaker is assigning greater prominence to the participants identified by the IC: as the game progresses the camera angle of the broadcast rapidly shifts from one player to another, focusing attention on who is in possession of the ball at every moment and thereby becomes the protagonist of the match (for further illustrations of how varying degrees of time pressure on the mental composition of utterances impacts upon their grammatical form see e.g. Mackenzie’s 2005 analysis of a commentary of a televised football game).

The infrequency of marked ICs seems to suggest that, as ICs are already “marked” constructions in terms of their formal and semantic configuration, as opposed to their unclipped counterparts, they offer comparatively more resistance than the latter to the “extra-markedness” imposed by preposing. For further details on the formal and discourse-functional features of fronting/preposing, the reader is referred to the corre-
sponding entries in e.g. Biber et al. (1999); Birner and Ward (1998); Dik (1978, 1997); Gómez-González (2001); Huddleston (1984); Lambrecht (1994); Quirk et al. (1985); Ward (1985); Ward, Birner and Huddleston (2002), and the references therein. Incidentally, the fact that just the EIF (without the RDC) may be placed before the it-predication – compare e.g. (12a) with *Gary Stevens who come forward it was, or A character defect it is that makes them go into politics it is – provides further evidence in support of our analysis that the EIF behaves as a discourse unit, on the one hand, and that it does not form a unitary constituent with the RDC, on the other. This, in turn, indicates that the RDC cannot be rankshifted or embedded within the EIF, but must be an immediate constituent of and therefore depends on the be-clause.

Lastly, only 34 tokens of interrogative it-clefts have been attested in ICE-GB (8.05%), and they tend to cluster around spoken discourse (30 cases (88.23%)/600,000, with a frequency of 0.5 per 10,000 words), as opposed to just a minority recorded in the written texts (4 instances (11.77%)/400,000, with a frequency of 0.1 per 10,000 words). The reason for this skewed distribution is the dialogic nature of the spoken ICE-GB texts, in which speakers are more likely to invite their addressees to take up the discourse floor through the use of questions.

Out of the 34 tokens, 22 are wh-questions (64.70%), and only 12 are yes-no questions (35.3%). Each of these two subtypes shows two patterns; accordingly four formal possibilities are obtained, which are illustrated in the examples in (13):

(13)  a. what is it that people in the country are concerned about (...) (ICE-GB:SIB-043 #122:1:B);
   b. what it is that’s going on (...) (ICE-GB:S1B-044 #122:3:A)
   c. (e.g. Was it you who was saying about uhm <,> moderate Shiites, (ICE-GB:S1A-042 #97:1:B)
   d. It was you that told me that wasn’t it, (ICE-GB:S1A-099 #272:2:A)

(13a) represents the most common interrogative IC in ICE-GB, showing the pattern wh-word(EIF) + be + it + RDC (12 tokens, representing 54.54% of the wh-subtype). It displays subject-verb inversion and mostly occurs in spoken texts (11 cases, 91.66%). Second in frequency is its non-inverted equivalent, that is, the wh-word(EIF) + it + be + RDC pattern, in (12b). There are 10 tokens of these (45.46% of the wh-type), and again they are mostly found in spoken texts (7 instances, 70%), where intonation alone can mark the interrogative pattern and resultingly there is no need for word order markers. The remaining 30% is found in written texts and all correspond to indirect questions which by definition cannot be inverted (e.g. what it is that frightens him, ICE-GB:W2B-023 #65).

Finally, the two variations of the yes-no subtype only appear in speech. Nine of them (75%) display the pattern be + it + EIF + DRC, as in (13c) above, while the other 3 are reduced versions of this pattern consisting of just be + it and deriving the other elements from the previous context, as in (13d), a negative question tag that appears after a positive statement.
4. Semantic features

ICs have been described as non-reversible identifying constructions of a grammaticalized (Huddleston 1984) or metaphorical (Borkin 1984) type as a result of the \textit{it} + \textit{be} predication that is encoded in the Head clause. This predication fossilizes the Identifier (IR) – Identified (ID) pattern onto the EIF and the dependent clause, respectively. Consequently, a special kind of discourse salience is conferred to the IR{EIF}, which usually is also the locus of intonational prominence. Following Borkin (1984:122), we would suggest that the core meaning of ICs is precisely to set up a relationship of identification of the specifying type in a given context, no matter what other communicative functions they may serve (see Section 5). This relationship of identification – as pointed out by Halliday (2004) – can be of three different kinds:

i. \textit{intensive}, the default case in which only a relation of identification of the specifying type is implied: this is the kind expressed in 78.67\% of our data (332 tokens) and has been illustrated in most examples considered so far, such as (1), (6a), (7a–e), or (8a–d);

ii. \textit{circumstantial}, when identification is combined with some kind of adverbial meaning and the relative constituent acts as Adjunct: we have seen that this relationship appears in 22.74\% of the \textit{it}-clefts in ICE-GB (96 items), mostly involving time (38.63\%), followed by place (35.22\%), manner (15.90\%), reason (7.95\%), purpose (1.15\%), and result (1.15\%), as shown in (14):

(14) a. Time: And it was only \textless \textgreater that the wound opened, ICE-GB:S1B-066 #25:1:A
b. Place: But it was in one of the office buildings that I discovered the letters \textless \textgreater, ICE-GB:S2B-023 #61:3:A
c. Manner: It was as a Guardsman that he came to the Second Battalion which now he commands and eventually became a lance sergeant instructor at the Guards Depot, ICE-GB:S2A-011 #124:1:A
d. Reason: It is for that reason that I have taken the assumption similar to that taken in many other areas of an oil price down to twenty-five dollars by the end of nineteen ninety-one \textless \textgreater, ICE-GB:S1B-052 #80:1:C
e. Purpose: it is only with the aim of achieving the same speed as the brain that new methods of implementation are sought, ICE-GB:W2A-032 #62:1

3. Several other terms have been used to contrast these structures. Thus, in transformational linguistics the distinction is between \textit{specificational} and \textit{predicational} structures (Akmajian 1979; Higgins 1979), whereas Kuno and Wongkhomthong (1981) speak about \textit{identificational} and \textit{characterizational} constructions. Halliday (1967a, b, 1994), in turn, call clefts predicated theme constructions and uses the term thematic equatives for what most people dub pseudo-clefts.
f. Result: it was as a result of something sticking out a piece of metal or whatever uh in the in the steps that’d caused that injury, ICE-GB:S1B-067 #156:1:A)

iii. and possessive, when both identification and possession are foregrounded: this is rather infrequent in ICE-GB (0.23%), and has the peculiarity that the marker of possession may be located either in the main clause, as in (15a), or in the dependent clause, as in (15b).

(15) a. It’s Mrs. Thatcher’s unselfish determination that free enterprise should not be held back by unwelcome explosive to public scrutiny that has allowed the industry to develop (ICE-GB:W2B-014 #6:1)
b. It’s Mrs. Thatcher whose unselfish determination (...) has allowed the industry to develop.

(15a) and (15b) above represent two presentational variants with slightly different processing effects. In (15a), both possessor (Mrs. Thatcher) and possessed (unselfish determination ...) are presented as IR/EIF/FA within one and the same attentional window. In other words, in (15a) the determination is in focus. By contrast, in (15b) possessor and possessed are presented in two separate moves: the possessor (Mrs. Thatcher) is in focus of the IC, while the possessed, though informative, is mapped onto the presupposed move of the construction as subsequent and subsidiary information.

In addition, it should be noted that the denotatum of the dependent clause is an open proposition, not a referential expression proper, in which information is presented as presupposed. In other words, the information borne by B corresponds to a presupposition that is true regardless of the truth value of the sentence as a whole. By way of illustration, let us reconsider the IC in (1a) above, now renumbered (16a). Here we would arguably have the open proposition in (16b) and the presupposition formally represented as (16c) (cf. e.g. Delin 1989; Lambrecht 2001; Prince 1978):

(16) a. It’s a character defect that makes them go into politics.
b. Something makes them go into politics.
c. ∃x makes them go into politics (x)

Both the presupposition and the open proposition contain an existentially-quantified variable (X) / ‘something’, whose identity is highlighted in (16a), a character defect, as a participant in the process of going into politics. This follows simply from the fact that ICs have a identifying (or “specification”) meaning: they specify a value for a variable. Thus (16) specifies the value a character defect for a variable in the open proposition “the X that makes them go into politics”, whose existence is presupposed. This suggests that the content of an IC is not exhausted by the statement of its presupposition, and also that it contains an assertion, to the effect that the EIF serves to instantiate the variable contained in the presupposition. An important function of ICs is therefore to separate presupposed and asserted material syntactically; and,
in addition, to indicate that an instantiation of the variable in the presupposition is taking place.

Before closing this section, mention should also be made of the implicature of exhaustiveness traditionally associated with ICs (with the exception of the there-subtype), which is often explicitly asserted by such words as only, precisely, just, actually, etc., as illustrated in (17) below (cf. e.g. Atlas & Levinson 1981; Davidse 2000:1121; Declerck 1984:271; Halliday 1967b:238; Halvorsen 1978:15; Hedberg 2000; Ward, Birner, & Huddleston 2002:1416):

(17) a. It needs <,> It does need the individual to set the whole <,> thing doesn’t it because it’s only the individual who can do it on (ICE-GB:S1A-045 #16:1:A)
    b. The only thing that I would say is that uh I’m not sure it’s just us <,> that think these atrocities are awful (ICE-GB:S1B-036 #31:1:D)
    c. The other thing is that it was actually Linda who said you can get a lot for a video in South Africa (ICE-GB:W1B-007 #11:1)

Broadly, the exhaustiveness feature of ICs has been attributed to the mappings of assertion, new information and intonational prominence onto the IR/EIF, on the one hand, and presupposition, given information and lack of intonational prominence onto the ID/RDC, on the other (in the unmarked instances). Compare the examples in (18), as opposed to John kissed Mary, where no exhaustiveness implicature nor existential presupposition is implied (adapted from Collins 1991:70):

(18) Assertion Implicature
    a. It was John that Mary kissed. positive positive
    b. It wasn’t John that Mary kissed. negative positive
    c. It was John that Mary didn’t kiss. positive negative
    d. It wasn’t John that Mary didn’t kiss. negative negative

Unlike Mary kissed John, the constructions in (18) all specify that out of the set of salient individuals that could have also been kissed by Mary (Peter, Tom, Paul, etc.) it was “John who was kissed by Mary”. These four constructions, however, differ in that those with identical assertions convey different implicatures, and those sharing the same implicatures have different assertions. Thus, both (18a) and (18b) share the implicature that “there is somebody that Mary kissed”, but while the former asserts that “Mary kissed John”, the latter asserts the opposite, namely that “Mary did not kiss John”. Similarly, both (18c) and (18d) presuppose that “there is somebody that Mary didn’t kiss”, but both assert opposite propositions: (18c) that “Mary didn’t kiss John” and (18d) that “Mary kissed John”.

In connection with this, there seems to be some disagreement as to whether or not “unicity” is an entailment of the exhaustiveness implicature of ICs. Thus, while some authors like Halliday (1967:236) seem to argue for it, others like Sornicola (1988:368–369) question this possibility arguing that “[i]t is doubtful whether the reply in (1)
[our (19)] [r]eally means that ‘It was John and only John’/’It was John and nobody else who came’.

(19) a. Who came?
   b. It was John.

Similar examples can be found in ICE-GB, in (20) below, in which unicity, or a “one EIF and only (one) EIF” reading, does not seem to be entailed either:

(20) a. It’s greed and self-interest or private interest that has caused the collapse (ICE-GB:S2A-057 #3:1:A)
   b. What was it that attracted you to Merlin (ICE-GB:S1B-048 #191:1:A)
   c. Whether it is a socially initiated kick that brings out technology, or technology that responds to an environmental kick is <unclear-word> particular to each example one may choose in relation to any society (ICE-GB:W1A-012 #69:2)
   d. However, it is not merely economic differences between the two countries that cause varying patterns of rural-urban migration (ICE-GB:W1A-013 #92:3)
   e. It is domestic politicians who communicate about the threat in order to mobilise public support for their own policy and power base. It is also the mass media who communicate about it in order to mobilise an increased audience (ICE-GB:W2A-017 #28-31:1)

In (20a), (20d), (20e) more than one EIF seem to be at issue, as implied by such additive markers as and, not merely, it is also. Likewise, in (20a) and (20c) more than one EIF are being considered: this time a disjunctive relation is established between two potential EIFs. Lastly in (20b) the it-cleft is formulated as an open question and therefore it is liable to receive an open answer which may, but need not satisfy a unicity interpretation.

A possible way out of the problem could be simply to claim that what is implicitly encoded in all it-clefts is an implicature of “exhaustiveness”, rather than one of “uniqueness”, the former being more generic or less specified than the latter (cf. Halvorsen 1978). It should also be kept in mind that implicatures, unlike entailments, can be cancelled or suspended in certain contexts, as it apparently happens in (19) and (20) above (cf. Horn 1981). Accordingly, it could be concluded that uniqueness is only entailed when made explicit in the context or when it is explicitly asserted by such words as only, precisely, just, and the like.

5. Discourse-cognitive features

In this section we shall first determine the informative dimension of the ICs in ICE-GB, in other words, we shall assess the informative status of its constituents against
the background of a context (Section 5.1). After this, we shall see how the informative patterns of ICs are directly related to their interactive dimension (Section 5.2), that is, their participation in the sequential and interactive organization of spoken and written discourse as topic (dis)continuity strategies (5.2.2) and as opinion-giving devices (Section 5.2.3), both side-effects of the thematic flexibility that characterizes these constructions (5.2.1).

5.1 Informative values

Before drawing on the informative values of ICs, a distinction should first be made between the relational and the referential interpretations of given and new information. For a fuller account of these notions and other associated categories the reader is referred to e.g. Gundel (1999, 2002) and Gómez-González (2001, 2004), and the references therein. For our purposes here it suffices to say that, while relational accounts are typically bipolar and analyze linguistic expressions in isolation (sentences, clauses or utterances), referential accounts work at discourse level taking into account not only the linguistic co(n)text, but also our model of the world and the speaker/hearer’s minds. Thus, in relational analyses the given represents information that is given in relation to the newsworthy part of a given message (sentence, clause, utterance, etc.) and vice versa, the new is more informative than the given. Some categories associated with relational givenness/newness are for instance such information-formal pairs as presupposition-focus (Lambrecht 1994), topic-comment (Gundel 1988) or topic-predicate (Erteschik-Shir 1997). By contrast, in referential analyses the given-new distinction is interpreted in terms of, for example, different degrees of assumed familiarity (e.g. Prince 1981a), different cognitive statuses of discourse referents (e.g. Chafe 1987, 1994; Gundel, Hedberg, & Zacharski 1993) or different kinds of recoverability (e.g. Collins 1991; Halliday 2004).

Here we shall interpret the given-new distinction referentially, more precisely, in terms of recoverability. Specifically, following Collins (1991) the label new will be used to refer to information that either has not been previously mentioned or is contrastive, that is, it establishes a contrast with respect to the members of a definite set that are contextually present. Given, on the other hand, designates information that has been previously mentioned and is expressed through such backgrounding strategies as anaphora, deixis, repetition, ellipsis, etc., or otherwise is indirectly retrievable by logical and cultural inference.

Now, turning to the informativity of ICs, we have already said that these constructions represent focal build-up devices because, through the predicative mechanism, the EIF becomes the FA and also tends to attract intonational prominence (Halliday 1967a, b, 2004; Hornby 1971). By placing, in the unmarked cases, a nuclear tone on the EIF itself, speakers provide this constituent with a contrastive flavor and, at the same time, emphasize that the identifying construction is incomplete until the ID/presupposition encoded in B is uttered. As a result, the EIF and the information encoded by it tend
to be presented and perceived as newsworthy, a tendency which explains why ICs have traditionally been said to be newness-oriented.

Although ICE-GB does not admit a reliable analysis of the intonational mappings of the given-new contrast as it offers no phonetic/intonation transcriptions of the texts, the referential exploitation of such a contrast and the three patterns obtained therefrom do seem to confirm the newness-orientation of ICs. As also attested in Collins (1991), these three patterns are: (i) new + given, (ii) given + new, and (iii) new + new (alternative or overlapping patterns have been proposed by e.g. Declerck 1984, 1994; Delin 1989; Hedberg 1990; Prince 1978).

Pattern 1, new + given, is displayed by ICs with an EIF that carries either non-recoverable or contrastive information, and a RDC that is given, as illustrated in the variations of this pattern collected in (21) below:

(21) a. new-given: Who communicates about the threat and for what purposes? The answer is extremely simple. It is domestic politicians who communicate about the threat in order to mobilise public support for their own policy and power base. It is also the mass media who communicate about it in order to mobilise an increased audience (ICE-GB:W2A-017 #28-31:1)

b. contrastive-given: If herd size is small and can't sustain a group alternative lifestyles must be resorted to. This leads to a completely sedentary way of life. However, as we can see from <unclear-character> Swift’s economic account of pastoralism, it is not just the poor who makes the transition to sedentism (ICE-GB:W1A-11 #24-26:1)

c. new-inferable: It’s not just not only the words is it they haven’t got the concepts either. Uh Is it is it Piaget who who who said about the different stage of development. You don’t understand abstract concepts until you’re much older, so they can’t express those properly <,> (ICE-GB:S1B-003 #79:1:D)

d. contrastive-inferable: There are too many books in this house. (…) Uhm I think I’m a bit too intense about books. So on going back to your to your childhood it was your mother wasn’t it who was the the driving force behind all of this behind this sort of intellectual rigour (ICE-GB:S1B-046 #54:1:B-57:1:A)

In (21a) above there are two ICs in a series; both have the same informative pattern with two new EIFs (domestic politicians and mass media) and two given RDCs directly recoverable from previous cotext (who communicate. . .). In (21b) the EIF the poor is contrastive with alternative group lifestyles, while the information encoded in the RDC is clearly recoverable from the repetition of completely sedentary way of life-sedentism: In (21c) the EIF, Piaget, is newly introduced, whereas the different stages of development in the RDC is inferentially recoverable via the simple deduction that the comprehension of words and concepts constitutes a gradual process. Finally, the EIF of (21d), your
mother, is contrastive with respect to the childhood figures that have been mentioned in the previous discourse (father, daughters) and could have influenced the speaker's intellectual rigour, while the information in the RDC is inferable via the association between books or the idea of being too intense about them and intellectual rigour.

Now turning to pattern 2, given + new, this applies to ICs which normally have a short EIF which carries typically anaphoric information and which prepares the context for the presentation of the RDC, which in turn encodes the more newsworthy part of the construction. Variations of this pattern are gathered in (22):

(22) a. given-new: In the Herbarium at Kew, botanists study plants from all over the world in order to identify and name them and to provide methods of helping other people recognise the plants around them. It is here that dried specimens end their journey. (ICE-GB:W2B-030 #8-9:1)

b. given-contrastive: The high salinity and wide temperature range are maintained because the Gulf has only a narrow connection to the Indian Ocean, which results in a limited exchange of water. And it is this last fact which makes marine pollution of any type so serious. (ICE-GB:W2B-029 #37-38:1)

c. inferrable-new: Ayatollah Khomeini, for instance, often killed as many of his own citizens in a single day as the Shah had killed in his entire reign. In the long term, it is the replacement of Arab dictatorships by democracies that will be the best guarantee of freedom, stability and peace in the Middle East (ICE-GB:W2E-001 #60-61:3)

d. inferrable-contrastive: Women have certainly not benefited from the greater economic success of other Third World countries to the same extent as men. ( . . . ) They have much lower access to education, to secondary and/or wage earning possibilities, in all countries across the Third World, although it is fair to say that it is in the least developed countries that they have suffered the most (ICE-GB:W1A-014 #30:1)

In (22a) above the EIF is a place deictic (here: the Herbarium at Kew) whose informational status is given; the RDC introduces the new idea of 'dried specimens ending their journey there'. The EIF in (22b) is given because anaphoric (this last fact), while in the RDC marine pollution is contrastively new, the contrast being with other kinds of pollution. In the case of (22c), the addressee is expected to draw an inference relationship between two dictators (Khomeini and the Shah) and dictatorships in order to decode the EIF (although neither of these gentlemen are Arabs, which could lead to the implication that only Khomeini was a dictator); whereas the RDC conveys new information, namely that the replacement of Arab dictatorships by democracies are a guarantee of freedom. The last example of pattern 2, in (22d) the addressee is expected to infer a link between countries across the Third World and least developed countries, while the RDC singles out the places where women suffer most (the least developed) out of the presupposed set (Third World countries).
In the last informational pattern of ICs, new + new, both the main and the dependent clause encode newsworthy information, as shown in the following combinations:

(23) a. new-new: It was in nineteen hundred and six that the Queen's great-grandfather King Edward the Seventh decreed that privates in the Household Cavalry should henceforth to be known as troopers (ICE-GB:S2A-011 #91:1:A)

b. new-contrastive: And I would like to say that it's the Labour Councillors on the Greater London Arts Board who are more to blame than anybody (ICE-GB:S1B-022 #105:1:C)

c. contrastive-new: I began this thesis with images of the positive, generous sexuality of the ancient Turkish deities and I conclude, with the inescapable fact that, bar the Tantric tradition, it is the sexuality of the Goddess, and consequently real women, that has suffered most in the process of transition that the Aryan heroes brought into the world's symbolic art forms so variously enshrined in each religious tradition (ICE-GB:W1A-008 #11:1)

d. contrastive-contrastive: In Kylie Tennant's Tiburon (1935), some women take part in a 1930s strike in which the unemployed refuse to 'scab' on other unskilled labourers, and therefore do not take jobs offered at a lower rate of pay than that agreed by the majority of such labourers. But it is the men who strike, and the women who dole out soup to the unemployed who are refused relief, that is, unemployment benefit (ICE-GB:W2B-009 #43-44:1)

In (23a) both the main and the dependent clause of this IC encode information that has not been mentioned before; whereas in (23b), only the be-clause conveys fresh news, while the RDC expresses contrastively new information: those that are *more to blame* is contrastive with respect to other factors mentioned in the previous and subsequent discourse which are co-responsible for “the present crisis in the arts in London” (ICE-GB:S1B-022 #60:1:D). In (23c), in turn, the EIF, the *sexuality of the Goddess* is contrastive with regard to the sexuality of male divinities as both sexes are implicitly present in the previously mentioned *generous sexuality of the ancient Turkish deities* (sexuality has by definition two values at least, female and male); while the content of the RDC has not been mentioned before. In (23d), on the other hand, both the EIFs (*labourers: men-women*) and the RDC contain contrastive information (*strike-dole out*).

The frequencies of the aforementioned patterns in ICE-GB are displayed in Table 4 below. The findings report pattern 1, new + given, as the most frequent, followed by patterns 3, new + new, and 2, given + new (51.65%, 29.62% and 18.72%, respectively), a tendency also reported in Collins (1991:111ff.) with the only difference that in the latter, pattern 2 outnumbers pattern 3.
Table 4. Informational classification of clefts in ICE-GB

<table>
<thead>
<tr>
<th></th>
<th>EIF</th>
<th>RDC</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNMARKED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a</td>
<td>New</td>
<td>Given</td>
<td>144</td>
<td>34.12</td>
</tr>
<tr>
<td>1b</td>
<td>Contrastive</td>
<td>Given</td>
<td>36</td>
<td>8.53</td>
</tr>
<tr>
<td>1c</td>
<td>New</td>
<td>Inferable</td>
<td>29</td>
<td>6.87</td>
</tr>
<tr>
<td>1d</td>
<td>Contrastive</td>
<td>Inferable</td>
<td>9</td>
<td>2.13</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td></td>
<td>218</td>
<td>51.65</td>
</tr>
<tr>
<td>MARKED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>Given</td>
<td>New</td>
<td>62</td>
<td>14.69</td>
</tr>
<tr>
<td>2b</td>
<td>Given</td>
<td>Contrastive</td>
<td>5</td>
<td>1.18</td>
</tr>
<tr>
<td>2c</td>
<td>Inferable</td>
<td>New</td>
<td>10</td>
<td>2.36</td>
</tr>
<tr>
<td>2d</td>
<td>Inferable</td>
<td>Contrastive</td>
<td>2</td>
<td>0.47</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td></td>
<td>79</td>
<td>18.72</td>
</tr>
<tr>
<td>3a</td>
<td>New</td>
<td>New</td>
<td>76</td>
<td>18.00</td>
</tr>
<tr>
<td>3b</td>
<td>New</td>
<td>Contrastive</td>
<td>13</td>
<td>3.08</td>
</tr>
<tr>
<td>3c</td>
<td>Contrastive</td>
<td>New</td>
<td>32</td>
<td>7.58</td>
</tr>
<tr>
<td>3d</td>
<td>Contrastive</td>
<td>Contrastive</td>
<td>4</td>
<td>0.94</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td></td>
<td>125</td>
<td>29.62</td>
</tr>
</tbody>
</table>

The fact that pattern 1 emerges as the unmarked one is probably a side-effect of the tendency of the global theme of ICs to encode in the EIF either non-recoverable information or contrastive information (as in (21) and (23) above). In the latter case the EIF establishes a contrast with respect to the members of a definite set, whether reinforced or not by contrastive focal stress. By contrast, the RDC tends to be associated with givenness because of its hypotactic status: on being subordinated to the impersonal *it + be* thematic predication, the RDC is imbued with a hedged, reportive flavor that presents it as logically presupposed, although it may also be informative, as illustrated in (22) and (23) above.

From the aforementioned the important observation follows that contrast is frequently, but not necessarily entailed in ICs, much in the same way as the RDC tends to contain referentially given information, but does not need to do so – as evidenced in the patterns just described – (cf. e.g. Borkin 1984:126; Collins 1991:155ff.).

5.2 The interactive dimension

In what follows we shall refer to the two main functionalities performed by ICs in discourse, their role as topic (dis)continuity markers (5.2.2) and their function as opinion-giving devices (5.2.3), both deriving from their characteristic thematic flexibility which is treated in turn.

5.2.1 Thematic flexibility

ICs admit two levels of thematic analysis: (i) a global or constructional pattern, in which the main clause (A) is the theme, or initial constituent, of the construction as a whole, and the dependent clause (B) is the rheme; and (ii) a local pattern, in
which both A and B have their corresponding theme-rheme arrays, that is, their corresponding clause initial participants, circumstances, processes or attributes (theme), and their rhematic continuations. As a result of the interaction of both patterns, in unmarked or prototypical cases, the EIF receives (at a local level) rhematic, syntactic and prosodic highlighting basically due to the principles of end focus and end weight; while (at a global level) the thematic predicative structure it is/was + EIF profiles the construction as a whole, (re)focusing the interactants’ attention on the EIF/IR which is affirmatively or correctively identified with the presupposed (ID) encoded in the subsequent or rhematic RDC (for a fuller account of this configuration see e.g. Gómez-González 2001).

Moreover, as attested through Section 3, the formal and functional variation exhibited by the EIFs of ICs demonstrates that these constructions allow for a thematic flexibility that is paradigmatic in orientation. In other words, ICs are capable of achieving maximum textual prominence by giving thematic, or initial, prominence and by putting in focus virtually any element that may occupy this slot. The preponderance of (pro)nominal subjects, followed by PPs as the most frequent EIFs in English can be explained in terms of the specific, identifying make-up of ICs. NPs have more referential potential than other categories, just as subjects and objects can be more easily construed as identifying than adjuncts. These in turn are more identifying than subjunctions, disjuncts and conjuncts which – unless metatextually – may never occur as the EIF. Furthermore, going by our data a proportional relationship could be posited between acceptability and factivity in such a way that the more factive-like the EIF and the RDC, the more acceptable the IC configuration seems to be. This tendency has also been hinted at by Borkin (1984:140), for example, when noting that the acceptability of some it-clefts with a clausal EIF improves if such expressions as the fact that are added (e.g. ??It’s that he shot her that he remembered vs. It’s the fact that he shot her that he remembered) (for a similar remark see also Ward, Birner, & Huddleston 2002:1418).

By means of the cleaving strategy subjects, the syntactic function which is least likely to carry intonational prominence in ordinary English declaratives, get formal and normally intonational prominence and at the same time are infused with discourse salience. Likewise, EIFs realized by adjuncts, direct objects, and the other syntactic functions, which are least likely to be thematic and as result least liable to profile the subsequent discourse, also receive through the cleaving strategy thematic and so additional discourse salience.

5.2.2 Topic (dis)continuity strategies

Endowed with the potential of this thematic flexibility, ICs reveal themselves as very versatile textual devices that mark topic (dis)continuity in discourse. Our data prove that they are used to focus attention upon a relation of topicality established by the speaker between the EIF with respect to the preceding and the subsequent discourse, an idea also emphasized by Halliday (1967:236) when explaining that in “It is Mary who ate the apple’, we will have the implicature ‘Mary and only Mary is the topic of the sentence’, and not ‘Mary and only Mary ate the apple’”.

Maria de los Ángeles Gómez-González
Specifically, ICs have been found to implement three main discourse strategies here dubbed topical, transitional, and corrective. Used in contexts of topic continuity, so-called topical ICs, typically displaying the given+new pattern, such as those in (22) above, are used to add information about the same topic as in the immediately preceding discourse segment(s). It should be borne in mind that topic continuity appears between two changes of discourse topic and it manifests itself because the same topic is kept at least between two adjacent discourse segments (a survey of the literature on topic (dis)continuity can be found in Goutsos 1977). Witness example (22a) now renumbered as (24) below:

(24) In the Herbarium at Kew, botanists study plants from all over the world in order to identify and name them and to provide methods of helping other people recognise the plants around them. It is here that dried specimens end their journey <.> (ICE-GB:W2B-030 #9:1)

This is a clear instance of a topical IC in which the EIF (here) points back to the previously established topic (the Herbarium at Kew) to introduce further information about it, namely the idea of ‘dried specimens ending their journey there’.

By contrast, transitional ICs, generally displaying the new + new pattern, such as those in (23) above, mark a (usually smooth) transition from one discourse stage to another by changing the topic or some circumstantial (usually spatio-temporal) aspect of the setting. A case in point is (23a) above now reproduced as (25) in a bigger context:

(25) And I’m sure he’d like to be <unclear-word> with them. Once again the guards present arms to the Sovereign’s Standard of the Life Guards. As the Life Guards trot past to the regimental trot The Keel Row or as we say in Scotland the The Keel Row <>. It was in nineteen hundred and six that the Queen’s great-grandfather King Edward the Seventh decreed that privates in the Household Cavalry should henceforth to be known as troopers. And uh <,> of course uh all troopers of the Household Cavalry on ceremonial occasions carry a drawn sword (ICE-GB:S2A-011 #91:1:A)

In (25) the IC indicates a smooth topic shift: the temporal EIF, in nineteen hundred and six, diverts attention from the time of the narration about Life Guards trotting past to The Keel to a flash-back scene in which we are informed about how privates in the Household Cavalry King Edward the Seventh came to be known as troopers, the topic of the subsequent discourse span.

The third type, corrective ICs, normally exhibiting the pattern new + given and therefore being the most frequent in our corpus, reformulate or correct (previous) topics, as in the examples in (20) above. The extract in (26) below provides a good illustration of this type, where the sports commentator uses three ICs almost in a series to “correct” attention towards what s/he judges to be the current topic of discourse, namely three different cyclists that lead the race at three different moments of the competition; but the idea of “who is the leader” is the global topic of the whole excerpt:
And the white helmet of the Bradford rider it is that takes up the lead. And disappointingly for Graham who was leading the first race when it was run he’s now at the back in third place and seemingly only heading for one point. At the moment it’s Paul Thorpe who’s heading for the maximum three points from this race. Sean Wilson a former Coventry rider now with Bradford he’s in second place at the moment but Thorpe is looking unassailable in the lead. Four laps of the circuit remember. They’ll go and take the yellow flag which will indicate now that they’re on the last lap. They take that yellow flag now. And it’s Thorpe who’s in the lead. In second place is Wilson in the blue helmet.

A remarkable feature of corrective \textit{it}-clefs is that in many cases they encode an explicit polarity contrast between what is or what is not the current topic of discourse as judged by the speaker, on the one hand, and the reversed polarity assumptions expressed in the previous and/or subsequent context, on the other. Thus, this kind of corrective ICs may display either a negative + affirmative pattern, as in (27), or an affirmative + negative sequence, as in (28); as the reader shall remember, most negative ICs in ICE-GB place the negative marker in the main clause. These two patterns, in turn, may be established between full clauses, as in (27a), or between the EIFs, in which case the pattern may be continuous, as in (27b), or discontinuous, as in (28), when the second EIF follows B, the dependent clause of an \textit{it}-cleft:

\begin{enumerate}
\item Still, it was not Rome that was merged in Italy, but Italy that was absorbed into the Roman citizen body. (ICE-GB:W2A-001 #22:1)
\item The borough electors, for example, knew nothing of the issues and as the voting turnout showed cared even less. It was not nominal electoral mass but political contacts which gave clout (ICE-GB:S2B-025 #32:1:A)
\end{enumerate}

And just as it is the development of this body of work that is the focus of my interest, and not Coleridge’s contribution to, or conformation with, distinct forms of intellectual practice, so this book is not a description of Coleridge’s return to Christian orthodoxy (ICE-GB:W2A-003 #10:1)

A final remark should be made that the aforementioned correlations between informative values and topic (dis)continuity-types of ICs represent only tendencies, not inviolable principles, as already noted with regard to their role as contrast markers or the given nature of the RDC in Section 5.1 above. For an exception to the correlations just described see, for example, (29) below:

\begin{enumerate}
\item It’s at this point that I think I can return to this theme of the name of the prophet (ICE-GB:S2A-036 #77:1:A)
\end{enumerate}

The IC in (29) above displays a given + new (deactivated) informative pattern, but it clearly marks a topic transition between two discourse stages, and therefore qualifies as a transitional IC, rather than as a topical one as would be expected from the general tendencies just described.
5.3 Opinion-giving devices

In addition and in close connection with the aforementioned, ICs also show an ability to indicate the speaker/writer’s attitude underpinning a proposition (Delin & Oberlander 1995). In this usage, they show varying degrees of involvement with respect to two opposite extremes, which Perzanowski and Gurney (1997:221–222) dub laudatory and accusatory ICs. Examples of each type are provided in the excerpts in (30) and (31) below:

(30) For many Conservatives, it is an article of personal and political faith that Britain must play a full part in European unification. During the long years of Labour's hostility to the EC, it was the Tories who kept the European flame burning in Britain. For the roles to be reversed would be a tragedy for many Conservative MPs and voters. (ICE-GB:W2E-004 #58:3-65:3)

(31) But Glenda Jackson politics is an essential part of this [the crisis of arts and entertainment industries in London] [. . .] Now you're a <,> Labour candidate <,> uh honest enough to say that there is a good deal of of party politics involved in this at the moment. Yes but it was a Conservative government that has systematically rate-capped so many Labour-led boroughs in and around London (ICE-GB:S1B-022 #53-58:1:D) [. . .] Now let John Mortimer come in there. [. . .] And I would like to say that it's the Labour councillors o o on the Greater London Arts Board who are who are more to blame than anybody for for cutting their subsidies to the arts (ICE-GB:S1B-022 #106:1:C)

In (30) the journalist is writing in favour of Britain’s integration into the European Community and uses an IC in a laudatory tone to stress the positive protagonism of the Conservative Party over the years in this process, as opposed to the ‘suicidal’ position against that process endorsed by Labour Party. By contrast, in (31), an excerpt from a broadcast discussion on the London arts, culture and entertainment industry, two ICs are used rather accusatorily by two different speakers with opposite views to put the blame for the situation on the two antagonistic parties: Glenda Jackson on the Conservative government and John Mortimer on the Labour councillors.

6. Distribution across ICE-GB text types

The interplay of the three characteristics of ICs described so far, i.e. formal, semantic and discourse-cognitive, explain their distribution across ICE-GB text types, displayed in Tables 5 and 6.

Tables 5 and 6 show that ICs are considerably more frequent in speech (286 tokens/600,000 words, occurring in a frequency of 4.8 per 10,000 words) than in writing (136 tokens/400,000 words with a frequency of 3.4 per 10,000 words). Interestingly
Table 5. Frequencies of *it*-decks across ICE-GB spoken texts

<table>
<thead>
<tr>
<th>Text category</th>
<th>No. of texts</th>
<th>Clefts</th>
<th>%</th>
<th>Subtotals Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIALOGUE S1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telephone calls</td>
<td>A: Private</td>
<td>45</td>
<td>4</td>
<td>1.40</td>
</tr>
<tr>
<td>Direct conversations</td>
<td>45</td>
<td>42</td>
<td>14.68</td>
<td></td>
</tr>
<tr>
<td>Classroom lessons</td>
<td>B: Public</td>
<td>20</td>
<td>18</td>
<td>6.30</td>
</tr>
<tr>
<td>Broadcast discussions</td>
<td>20</td>
<td>21</td>
<td>7.34</td>
<td>127</td>
</tr>
<tr>
<td>Broadcast interviews</td>
<td>10</td>
<td>13</td>
<td>4.54</td>
<td>155%</td>
</tr>
<tr>
<td>Parliamentary debates</td>
<td>10</td>
<td>11</td>
<td>3.84</td>
<td>2.1</td>
</tr>
<tr>
<td>Legal-cross examinations</td>
<td>10</td>
<td>16</td>
<td>5.60</td>
<td></td>
</tr>
<tr>
<td>Business transactions</td>
<td>10</td>
<td>2</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>MONOLOGUE S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spontaneous commentaries A: Unscripted</td>
<td>20</td>
<td>53</td>
<td>18.53</td>
<td></td>
</tr>
<tr>
<td>Unscripted speeches</td>
<td>30</td>
<td>36</td>
<td>12.58</td>
<td>150</td>
</tr>
<tr>
<td>Demonstrations</td>
<td>10</td>
<td>8</td>
<td>2.80</td>
<td>52.45%</td>
</tr>
<tr>
<td>Legal presentations</td>
<td>10</td>
<td>8</td>
<td>2.80</td>
<td>2.5</td>
</tr>
<tr>
<td>Broadcast talks B: Scripted</td>
<td>20</td>
<td>25</td>
<td>8.74</td>
<td></td>
</tr>
<tr>
<td>Non-broadcast speeches</td>
<td>10</td>
<td>20</td>
<td>7.00</td>
<td></td>
</tr>
<tr>
<td>MIXED S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broadcast news</td>
<td>20</td>
<td>9</td>
<td>3.15</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td>TOTAL</td>
<td>300</td>
<td>286</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Frequencies per 10,000 words</td>
<td>600,000 words</td>
<td>4.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>67.77%</td>
</tr>
</tbody>
</table>
Table 6. Frequencies of *it*-clefs across ICE-GB written texts

<table>
<thead>
<tr>
<th>Text category</th>
<th>No. of texts</th>
<th>Clefts</th>
<th>%</th>
<th>Subtotals</th>
<th>Percentages Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>NON PRINTED W1</td>
<td>Untimed student essays A: Non professional 19.85%</td>
<td>10</td>
<td>13</td>
<td>9.55</td>
<td>10 40</td>
</tr>
<tr>
<td></td>
<td>Student examinations scripts</td>
<td>10</td>
<td>14</td>
<td>10.30</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Social letters B: Correspondence 9.56%</td>
<td>15</td>
<td>8</td>
<td>5.88</td>
<td>29.41</td>
</tr>
<tr>
<td></td>
<td>Business letters</td>
<td>15</td>
<td>5</td>
<td>3.68</td>
<td>1.0</td>
</tr>
<tr>
<td>PRINTED W2</td>
<td>Humanities A: Academic 26.48%</td>
<td>10</td>
<td>17</td>
<td>12.50</td>
<td>10 29.41</td>
</tr>
<tr>
<td></td>
<td>Social sciences</td>
<td>10</td>
<td>7</td>
<td>5.15</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Natural sciences</td>
<td>10</td>
<td>5</td>
<td>3.68</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Technology</td>
<td>10</td>
<td>7</td>
<td>5.15</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Humanities B: Non-academic 30.88%</td>
<td>10</td>
<td>9</td>
<td>6.62</td>
<td>10 70.59</td>
</tr>
<tr>
<td></td>
<td>Social sciences</td>
<td>10</td>
<td>10</td>
<td>7.35</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Natural sciences</td>
<td>10</td>
<td>12</td>
<td>8.82</td>
<td>70.59</td>
</tr>
<tr>
<td></td>
<td>Technology</td>
<td>10</td>
<td>5</td>
<td>3.68</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>Press news reports C: Press</td>
<td>20</td>
<td>6</td>
<td>4.41</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Administrative D: Institutional</td>
<td>10</td>
<td>1</td>
<td>0.74</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Skills/hobbies</td>
<td>10</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Persuasive writing E: Editorials</td>
<td>10</td>
<td>4</td>
<td>2.94</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>Creative writing F: Fiction</td>
<td>20</td>
<td>13</td>
<td>9.55</td>
<td>1.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>200</td>
<td>136</td>
<td>100</td>
<td>100</td>
<td>32.22</td>
</tr>
</tbody>
</table>

Frequencies per 10,000 words 400,000 words 3.4
enough, these findings differ from other investigations. Collins (1991), for example, reaches the opposite conclusion: ICs in the written mode outnumber those in the spoken mode, though only by a narrow margin (565 examples, with a frequency of 5.7 tokens per 10,000 vs. 187 examples occurring in a frequency of 4.3 tokens per 10,000 words). Similarly, Miller and Weinert (1998:197) explain that it-clefts “are a regular highlighting device in written English but are rare in my spoken data...”. Given the relative recency of our data, it could be tempting to suggest that there has been a change in the usage of ICs throughout time so that they have been transferred from the written to the spoken mode. Yet, considering the limited size of our data, this hypothesis still awaits confirmation of more exhaustive analyses tracing the formal and functional evolution of the construction through time.

The properties of ICs already described, i.e. their specificational meaning, their formal flexibility and their newness-orientation, resulting in their usage as topic (dis)continuity strategies and/or as an opinion-giving devices, make these constructions particularly suitable for spontaneous and argumentative discourse. This is precisely the rhetorical function of the spoken ICE-GB texts they mostly cluster around: (i) spontaneous commentaries (18.53%); (ii) direct or face-to-face conversations (14.68%), (iii) unscripted speeches (12.58%), (iv) broadcast talks (8.74%), and broadcast discussions (7.34%). In all these textual categories it is natural to find spontaneous or argumentative discourse aiming to express one’s perspective or opinion on the experienced being construed. ICs are suitable tools for these (sub)genres in that they simultaneously express specificational meaning and associate typically new or contrastive information with the theme/EIF, thereby (re)directing the addressee(s)’ FA towards the speaker’s topic and his/her perception of the ongoing discourse.

Indeed, although apparently more frequent in monologic speech (150 tokens (52.45)/600,000, 2.5 frequency) than in dialogic speech (127 tokens (44.40%), ICs are appropriate tools to collaborate on the negotiation of the discourse field, a characteristic of both speech modes. For the communicative act to be successful, the discourses created by the different participants have to share the same fields in the case of dialogue and, with regard to monologues, both speakers and addressees contribute to the negotiation of the domain. The interpersonal flavour of ICs and their ability to mark topic (dis)continuity make them well suited for this negotiating process. This is particularly evident in broadcast discussions in which speakers use ICs to specify their viewpoints on the issues discussed (as in (30) and (31) above). Similarly, in sports commentaries, where the audience’s perception of reality is mediated by the “camera movements” of the broadcaster, commentators use ICs to specify the FA as reality unfolds (as in (12), (21), (26) above).

In writing, by contrast, the sense of institutional focus must be achieved by the writer. The lack of simultaneity in the creation and perception of the written message prevents the reader from negotiating the field; s/he has to discover it instead. In these contexts ICs become useful devices to help the reader achieve this process of discovery: by specifically identifying the FA they (re)direct the reader’s attention towards a particular direction following the writer’s dictates. This guiding purpose pervades
the descriptive or informative written ICE-GB texts ICs cluster around, namely Humanities (12.50%), student examination scripts (10.33%), creative writing (9.55%), untimed student essays (9.55%), and (non-academic) Natural Science (8.82%).

Last, but not least, our data confirm that there exists a direct proportional relationship between the frequency of ICs and the degree of formality of the communicative situation, that is, the extent to which discourse is preplanned or not. The findings reveal that ICs are typically associated with conditions of informality: more precisely, with spontaneous discourse in the spoken mode (e.g. direct conversations, commentaries, unscripted speeches, broadcast discussions); and with non-academic and non-professional writing (within the printed and non-printed categories, respectively).

7. Conclusion

Although a number of features such as contractiveness, focus, and exhaustiveness have been commonly identified as intrinsic features of ICs, here it has been argued that the core characteristic of this construction is rather its identifying-specificational meaning. This derives from the constructional pattern “EIF, or some aspect of it, is identified with a variable (X) in B, which encodes an open proposition”, or put another way, “the variable (X) in the open proposition in B equals the EIF”, which in turn becomes the FA of discourse. Furthermore, the specific identification of ICs has been adduced to explain why the EIF tends to be referential and specific (rather than characterizing). Specifically, our findings have reported the it is/it was … that as the most recurrent type, usually with a subject/agent NP in the EIF slot.

In addition, it has been explained that, although they are suitable for expressing contrast, the EIF of ICs is not always in contrast, much in the same way as the information encoded in the dependent clause need not be always given. The pattern new + given has been qualified as the most frequent, followed by the patterns new + new, and given + new. These three configurations have been related to three main topic (dis)continuity strategies: (i) topical, when ICs, generally with the pattern given + new, take up and push forward a previous topic; (ii) transitional, in those instances in which ICs (re-)introduce new or deactivated topics or spatio-temporal settings, normally exhibiting the new-new values; and (iii) corrective, or ICs that reformulate or correct old/previous topics, with frequently the new + given pattern. But, again, these have been regarded merely as tendencies, not as inviolable principles.

Lastly, we have seen that ICs are useful tools for expressing the speaker’s/writer’s view of, or involvement towards, the content presented in the proposition, which, in combination with the aforementioned, explains their tendency to cluster around informal or spontaneous opinion-giving spoken texts.
References


International Corpus of English (ICE) http://www.ucl.ac.uk/english-usage/ice


The Survey of English Usage (ICE) http://www.ucl.ac.uk/english-usage


Another take on the notion Subject*

Dik Bakker and Anna Siewierska
University of Amsterdam / University of Lancaster

Linguistic theories typically have a static, competence-based view of the notion Subject. In this view speakers choose their Subjects freely from the relevant constituents of a clause, typically the arguments of the main predicate. However, when one looks at what speakers actually do, it becomes clear that the choice is determined to a high degree on the basis of the pragmatic and semantic features of the candidate constituents. Apart from roles such as Agency and Topicality, the major determining factor appears to be a possibly language specific mix of features such as person, animacy, and definiteness. A multifunctional approach to Subject choice seems to work even for English, a language which is thought to have grammaticalized Subject to a large extent.

1. Introduction

Subject is both one of the most controversial notions in linguistics and the one most often taken for granted. In traditional linguistics it is typically inherited from Latin grammar, and is seen as the constituent that has Nominative case and agreement on the verb. More fundamentally, it is seen as the most central constituent of the clause, one which also plays a role in more complex syntactic constellations, such as equi NP deletion. In the various versions of generative theory, Subject is typically treated as a ‘deep’ syntactic position, on which different constituents may land via the application of transformational rules. Relational Grammar (Perlmutter 1982) shares this double deep-to-surface aspect of grammatical relations, i.e. Subject and Object, be it that in this case there are links with semantic and pragmatic aspects of the structure of the clause.

* These ideas were first presented at the International Symposium on the Typology of Argument Structure and Grammatical Relations in Kazan, May 2004 (Siewierska & Bakker 2004). We wish to thank Nick Smith from the University of Lancaster for providing us with the corpus data discussed in Section 2. We also wish to thank the editors and an anonymous reviewer for suggestions which led to important changes in the original version.
As we will see in more detail below, functional approaches to linguistic theory, and more specifically Functional Grammar (FG; Dik 1997), Role and Reference Grammar (RRG; Van Valin & LaPolla 1997) and Relational Typology (RT; Kibrik 1997), take slightly different positions in the area of grammatical relations. What they seem to share, however, is that Subject is seen as a surface syntactic function correlating to but assigned independently from underlying pragmatic and semantic functions and features. Under the guise of notions such as ‘Perspective’ (FG), ‘Priviliged Syntactic Argument’ (RRG) and ‘Pivot’ (RT), these theories define Subject indeed as a self-standing syntactic function assigned by the speaker to the most central constituent of a clause. It is chosen more or less independently from other semantic and pragmatic features of the constituents to which the grammar in question limits the choice. Mutatis mutandis, this state of affairs is not completely unlike the position of Subject in formal syntax.

We think that this view of a more or less freely assigned Subject appears to work only for a restricted number of languages, mainly from Western Europe, West Africa and New Guinea. It is almost exclusively in these linguistic areas that languages may be found which have passives with the presence of Agents, albeit optional and typically in the form of an oblique. Such a constellation, with all arguments present, may in turn be interpreted as the result of a grammar-induced choice between the arguments for the central position. Outside the linguistic areas mentioned, passives tend to be obligatorily agentless. However, most of the other morphosyntactic phenomena traditionally related to Subjects are to be found in all areas, and are therefore in need of an explanation as well.

In this contribution, we will propose a less grammaticalized and more dynamic take on the notion Subject, not unlike Givón’s (1997) multifactor approach, which we think may be applied to the vast majority of the languages of the world, including those mentioned earlier on. We will go about our job in the following fashion. First, in Section 2, we will try to show that even in a grammar of English, a language which is often seen as having grammaticalized Subject to the extreme, there are other factors that determine which of the constituents in a clause is going to be Subject and that it is not just a free choice between the arguments by the speaker. We will present our observations in terms of one of the leading functional theories to date, Functional Grammar (cf. Dik 1997). Within this theory, Mackenzie (1998) is one of the first contributions drawing attention to the study of sentences in discourse in a principled fashion. Our conclusions, however, are general and might be embedded in any functionally oriented theory. In Section 3 we will sketch the outlines of a dynamic, multifunctional framework for the description of Subject-like phenomena which has precisely this aim. Section 4 will discuss some complex and problematic cases, and in Section 5 we will draw some final conclusions.
2. Subject in a grammar of English

In general terms the Subject in an English sentence is the constituent involved in a number of morphosyntactic phenomena. Following Keenan (1976), we divide them into the coding and behavioural properties of Subjects. Under coding properties we find case marking (typically Nominative or Absolutive for Subjects; Accusative and Ergative for Objects) and agreement marking on the verb (typically, the marker varies for Person, Number and Gender features of the Subject constituent). Behavioural properties are syntactic phenomena under the control of Subjects. Examples include anaphoric binding as expressed in reflexive pronouns, conjunction reduction, equi NP deletion, raising, and the relativized constituent in relative clauses, among other things. Finally there is constituent order, which may be both under the control of Subjects and employed as a device to mark them. In Indo-European languages it is often the case that all these phenomena, to the extent that they are relevant for a language in the first place, are related to one and the same constituent. English is a case in point. So, in example (1a) she has Nominative marking and controls verb agreement and the empty Subject position in the coordinated clause. Example (1b) shows that the passive construction may reverse the situation, promoting the Patient argument to the Subject position, where it takes over all the coding and behavioural properties of the Agent in (1a).

(1) a. She
nom
approaches
Jack
acc
and Ø
kisses
himi
/* Ø
kisses
her

b. Jack
nom
is
approached by her
acc
and *Ø
kisses
himi
/* Ø
kisses
her

In a rendition of the grammar of English which is completely focussed on linguistic competence, it could be stated that Subject may be assigned to Agents and Patients, and also to Recipients, but not to for instance Instruments or Locatives, as shown in (2a–c).

(2) a. She
rec
was sent a Valentine by a one of her pupils

b. *The hammer
instr
was opened the door by him

c. *Home
loc
is stayed by her

Functional Grammar (Dik 1997:267) proposes a language-dependent but universally valid Semantic Function Hierarchy for the assignment of Subjects. It takes the following shape:

(3) Semantic Function Hierarchy (SFH):
   Agent > Patient > Recipient > Beneficiary > Instrument > Locative > Temporal

For each language there is a cut-off point. All semantic functions to the left of this point may have Subject assigned to them; those on the right may not. For English the cut-off point on the SFH is between Recipient and Beneficiary. In other languages it is located more to the left (e.g. for Dutch after Patient) or to the right (e.g. for Sundanese after Beneficiary and for Malagasy after Locative). Thus, Subject may – and can – be
assigned to any constituent with a semantic function to the left of the cut-off point, irrespective of its further semantic and pragmatic properties, or other aspects of the clause in question. The decisive criterion for the choice of Subject, according to FG, is an independent principle called *Perspective*. In other words, speakers choose their subjects within the relevant limits of the SFH in order to present the contents of the clause from the perspective of the referent in question.

Such a rendition of the Subject assignment rules does not, however, seem to reflect linguistic reality; in other words, it does not comply with linguistic performance, and how speakers actually operate. Compare the acceptability of the sentences in (4a–d).

(4)  
- a. She bought a new bike.
- b. ?A new bike was bought by her
- c. ?A falling stone hit her
- d. She was hit by a falling stone

Although all four versions might be considered well-formed in terms of competence-based rules, only those with pronominal subjects, (4a) and (4d), seem acceptable from a performance point of view in the sense that they might be expected to appear in normal discourse rather than as isolated examples in a linguistic discussion of English. In fact, utterances such as (4b) and (4c) are very rare in corpora of spoken English; (4a) and (4d) seem to be the natural, and in fact the only options. A search through the spoken section of the British National Corpus (BNC; Aston & Burnard 1998) confirms that speakers select their Subjects on the basis of their semantic, pragmatic and discourse properties rather than on the basis of a more or less independent operation such as *Perspective*. Firstly, passives are rare in the spoken language. Moreover, passives with explicit agents are rarer still. Svartvik (1966), in a corpus of written English, found that over 80% of the passives were agentless. For a corpus consisting of a mixture of written and (formal) spoken English, Thompson (1986) found almost exactly the same. The nature of the corpus selection process did not allow us to inspect all candidate sentences for agentless passives. Therefore, we can only more or less impressionistically state that passives with agents are just a fraction of all passives to be found in the spoken corpus, probably considerably less than 15% of them. It may be concluded, then, that the major reason for speakers of English to opt for the passive is the absence of an agent since it is unknown, irrelevant or obvious, i.e. for discourse pragmatic reasons. According to Thompson (1986) it is also for discourse pragmatic reasons that speakers of English opt for a passive despite the presence of an agent. The non-agent (i.e. the second argument) will be made Subject when “(it) is more closely related than the agent to either the theme of the paragraph or to a participant in the immediate preceding clause”. Obviously, this may be interpreted in terms of the different types of discourse and sentence topics.

1. Givón (1979:59) assumes that missing agents are always recoverable by the hearer, and therefore left out by the speaker, at least in terms of the type of referent.
Another take on the notion Subject

Table 1. Distribution of types of passive Subjects and Agents in the BNC

<table>
<thead>
<tr>
<th>Subject</th>
<th>Agent ('by')</th>
<th>number of occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronominal</td>
<td>Pronominal</td>
<td>14 (8.5%)</td>
</tr>
<tr>
<td>Pronominal</td>
<td>Nominal</td>
<td>102 (61.8%)</td>
</tr>
<tr>
<td>Nominal</td>
<td>Pronominal</td>
<td>2 (1.6%)</td>
</tr>
<tr>
<td>Nominal</td>
<td>Nominal</td>
<td>47 (28.5%)</td>
</tr>
</tbody>
</table>

That topicality, and pragmatics in general, is an important parameter in the determination of Subjects in English may be gathered from the following data from the BNC. In the section with spoken English we found 236 utterances which contained both a form of be and the preposition by. Of these, 71 were analysed as non-passives or in fact agentless passives. The 165 agented passives that remained may be characterized as follows in terms of the form of the two arguments.

The Subject is thus a pronoun in over 70% of the cases, and the Oblique Agent is pronominal in only 10% of the cases, which is indicative of the overall topicality of the referent in the Subject position. In almost two thirds of the cases, we find the combination of a pronominal Subject and a nominal Agent. When both referents are pronominal all cases but one follow the person hierarchy in (5) below.

(5)  *Person hierarchy*

1st > 2nd > 3rd

The only counterexample has a relative pronoun for its 3rd person subject, so that there is no real choice at all. The two sentences which have a nominal subject and a pronominal Agent can be found in (6) and (7) below.

(6)  and will continue to, to, to, erm so that the, the, the set-up should never have been created by hér. [PS527]

(7)  We said that come from the film Cats and he reckoned the copy of Memories he’s got is sung by who? [PS0FX]

In both cases the Subject is topical while the Agent is focal, and therefore prefers sentence-final position. Finally, of the 47 nominal pairs, the majority have a definite Subject and an indefinite Agent. However, in eight cases the Subject is indefinite while the Agent is definite. Of these, six have a human Subject while the Agent is non-human. The other two are given in (8) and (9):

(8)  On their births a trust a fund of tens of thousands of pound was started for them in their names by their grandmother. This will guarantee all their school fees and the basis of their future. [PS0FG]

(9)  And the planet earth is within the reach of gravitational pull and things on the earth are attracted by the moon. [KPAPS000]

In (8), the money is going to be the new topic for the following stretch of discourse. In (9), the earth has been introduced as a topic, and is continued by the related sub-topic
in the Subject position of the second clause, while the moon – semantic function Force, not Agent – is in focus.

From the above discussion, and from earlier work by authors such as Thompson (1986), a picture of Subject in English emerges which is much more embedded in the dynamics of discourse than the purely sentence-grammar-oriented treatment in standard Functional Grammar and other theories. In the next section we will try and sketch a more encompassing framework for the treatment of Subject in English, and languages in general.

3. Towards an encompassing framework for Subject

The ingredients for our general approach to Subject stem from two domains of linguistic description: the functional and the formal one. Under the functional domain we subsume all relevant semantic and pragmatic features, such as the respective semantic and pragmatic functions, and notions like animacy, definiteness and number. In FG, they pertain to the underlying representation of a clause, and are typically aspects of referring elements, such as argument and adjunct positions. Although their role may be different for different languages, they are assumed to be drawn from a universal set and to be subject to the respective universal hierarchies that have been proposed in the typological literature, as illustrated in (3) and (5) above by the semantic function and person hierarchies. Under the formal domain we subsume all morphosyntactic phenomena which are controlled by an argument or adjunct, and which were subdivided into coding, behaviour and constituent order in Section 1. Also in this case, we assume that a maximum set may be determined from which only a subset may be relevant for a specific language. The Subject-related morphosyntactic phenomena (MSP) are the explananda, whereas the explanation rests on the relevant set of pragmatic and semantic factors (PSF). In short:

\[
\text{PSF}_k \rightarrow \text{MSP}_k
\]

where PSF$_k$ and MSP$_k$ are the relevant subsets for language k, and the arrow should be read as 'control'.

We suggest the following interpretation of Subject. Arguably, all languages have MSP which are controlled via one or more arguments or adjuncts. When for any given language it can be shown that at least two basic semantic functions, such as Agent and Patient, can each control the relevant set of MSP, the control function will be seen not as semantic but syntactic. Subject is the label belonging to this unified syntactic function. We will say that the relevant semantic functions are neutralized to Subject. In order for this neutralization to hold in some language, and for Subject to be relevant in the first place, it should be shown on independent grounds that the semantic functions in question play an independent role elsewhere in the grammar of that language. For example, they may have different case markings when they are non-Subjects, trigger
Another take on the notion Subject

different grammar rules, etcetera. Thus, looking at the examples in (11) below from Acehnese, spoken in Indonesia, we would not assume Subject to be relevant for this language.

(11) Acehnese (Austronesian; Durie 1985)
   a. Gopnyan geu-tém jak/*geu-jak.
      3sg 3-want go/*3-go
      ‘He wants to go.’
   b. Gopnyan geu-tém *hët/geu-hët.
      3sg 3-want *fall/3-fall
      ‘He wants to fall.’

In (11a) the subordinate clause has obligatory equi-deletion of its (single) argument, thus rendering the clause infinite. In (11b), however, equi-deletion is forbidden: the argument is obligatorily present, resulting in a finite subordinate clause. The decision on the expression of the argument hinges crucially on the semantic function of the argument of the embedded verb: Agents get no expression, and Patients are expressed via a clitic. Since this agreement rule is based on individual semantic functions rather than on a neutralized set of them, there seems to be no reason to posit a Subject function for Acehnese. We will call such a language role-dominated.

On the other hand, there are languages in which it seems to be pragmatic functions which control some of the MSP under discussion. The example in (12) below, from Mongsen, a dialect of Naga Ao, is a case in point. Topical constituents undergo conjunction reduction, irrespective of their semantic function. Such a language does not seem to be in need of an independent Subject function either.

(12) Mongsen (Tibeto-Birman; Coupe 2003)
   t(e) ai-t(e)-p(e)ti a hl(e)p-(e)r
   thus NRP-be.big one wrap-seq
   n(e)-ni m(e)-thu? t(e)si a? pi t(e)-s(e)la
   2sg.POSS NEG-reach until just PROX PROX-untie
   t(e)-(e)r aj ok h(e)m(e)t-(e)r ai oj z(e)k
   thus-seq PRT PRT hold-seq PRT PRT send.PAST
   ‘Thus having wrapped a big one [a bundle], [Lichabai said to Mechatseg]
   “Don’t untie [this] before you reach your wife”. And then, taking hold [of it, i.e. the bundle], [Lichaba] sent [him].’

---

2. Here, we follow Dikker (2004), who argues for establishing the set of relevant semantic functions in a language on language internal grounds rather than assuming that the set of semantic functions is universal and given a priori for any language. A test for relevance of a tuple of semantic functions is particularly crucial for functions which typically share the same argument position, and are therefore mutually exclusive, such as Agent and Force.

3. RRG (Van Valin 1996:274f.) would call Acehnese a language with a semantic rather than a syntactic pivot.
In the first line of this short stretch of discourse, a bundle is introduced with Patient function. In the second line explicit mention of it is left out again from the Patient position. In the third line, all three topical referents are left out from the respective Agent and Patient positions. Languages such as Mongsen will be called pragmatically dominated.

However, in most languages the state of affairs seems to be less clear cut. The relevant MSP are neither fully role-dominated nor fully pragmatically dominated, but controlled by a set of semantic and pragmatic properties of the clause or sentence. In these languages control over MSP is assumed to be syntactic, and Subject is taken to be the function assigned to the argument or adjunct which exerts this control. The semantic and pragmatic factors concerned could be interpreted both as motivating and as constraining factors for Subject assignment in a specific language. We distinguish the following four subdomains for these factors:

a. Semantic functions
b. Pragmatic functions
c. Semantic and formal features of relevant argument and adjunct positions
d. Other pragmatic, semantic and formal aspects from the context

First, there is the set of semantic functions which can be neutralized under Subject assignment. They may be found on the Semantic Function Hierarchy as given under (3) above. As indicated earlier, we take into consideration individual semantic functions rather than argument positions, as in Functional Grammar. Therefore, the following two examples suffice for a language to have Subject. In example (13a) below, control over the equi-deleted argument from the second conjunct is exerted by the Agent intransitive argument of the first conjunct, while in (13b) the same is done by a Patient intransitive argument.

(13) Dutch

a. Vader sprong en viel
   father\textsubscript{AG} jump\textsubscript{PAST} and Ø\textsubscript{PAT} fall\textsubscript{PAST}
   ‘Father jumped and fell.’

b. Vader viel en stond weer op
   father\textsubscript{PAT} fall\textsubscript{PAST} and Ø\textsubscript{AG} stand\textsubscript{PAST} again up
   ‘Father fell and stood up again.’

Obviously, Subjects are relevant for all languages where neutralization takes place over both arguments of a transitive predicate, typically accompanied by a passive or antipassive construction, as in the examples in (14) from Dyirbal.

(14) Dyirbal (Australia; Dixon 1994:13f.)

a. Y abu nguma-nggu bural-n banaga-n\textsuperscript{u}
   mother\textsubscript{ABS} father\textsubscript{-ERG} see\textsubscript{-NONFUT} return\textsubscript{-NONFUT}
   ‘Father saw mother and she returned.’
Another take on the notion Subject

b. Nguma bural-nga-n’u yabu-gu banaga-n’u
father.abs see-antipass-nonfut mother-dat return-nonfut
‘Father saw mother and returned.’

In (14a) the Patient argument is zero marked, is located in sentence-initial position, and controls conjunction reduction. In (14b) the same holds for the Agent argument.

The second factor are pragmatic functions. Even in languages which are not fully pragmatically dominated, pragmatic functions play a central role in the assignment of Subjects. Indeed, Topics are so central in the organization of sentence structure that Subject is often seen as grammaticalized topic (cf. Givón 1983). It is quite unlikely that Topic will not turn out to be a more or less central element for the assignment of Subjects in the majority of languages, as shown for the English corpus examples in Section 2. Focus, by contrast, seems to a less determining factor for Subjecthood crosslinguistically. Although focal constituents, such as Q-words, may be fronted in many languages, this is typically either just a marked word order variant, or a feature of a special construction such as clefting. The other MSP typically do not apply, as in the case of passivization.

The third factor are the semantic, and arguably also formal, properties of the relevant argument and adjunct positions. In the English corpus examples, we saw that the person hierarchy under (5) was instrumental in promoting 1st and 2nd person referents – pronominal almost by definition – to Subject position at the cost of 3rd person referents. In the case of two nominal arguments, definiteness and humanness turned out to be decisive factors. These factors play a role in many languages. Sometimes they seem to be grammaticalized to the extent that the alternative is ungrammatical. This is shown in example (15) from Navajo. Another manifestation of the role of such factors is the occurrence of so-called inverse constructions in languages like Nocte, as shown in example (16).

(15) *Navajo* (Athapascan; Witherspoon 1980:5)

a. At’ééd to yo-odlāq
   girl water act-drink.past
   ‘The girl drank the water.’

b. *To* at’ééd bo-odlāq
   girl water pass-drink.past
   ‘The water was drunk by the girl.’

---

In some languages the form referring to Speaker and Addressee may be of a nominal nature. In Thai, for instance, speakers use the word *phom* ‘hair’ in self-reference when speaking to a superior, and *tua* ‘body’ for the Addressee when this is someone closely related, e.g. a spouse (cf. Siewierska 2004:228). It remains to be seen whether these behave like pronouns or like nouns in the corresponding languages.
In Navajo, non-animates cannot be Subjects when there is an animate alternative, as in the case of (15b). In (16a) we have agreement with the first person Agent. In (16b), where the first person is found in the Patient slot, we nevertheless have agreement with the first person, be it that the verb is marked for non-default Subjecthood.

In this third category we may mention as a factor the nominal versus pronominal status of the referent. We saw that the vast majority of English Subjects in our BNC sample are pronominal. In Cree, the use of the inverse seems to be based on the pronominal vs. nominal form of the arguments of transitives, at least to some extent. Compare examples (17a) and (17b).

(17) Cree (Algonquian; Wolfart 1973)

a. Ni-sekih-a atim.  
I-scare-dir dog  
‘I scare the dog.’

b. Ni-sekih-iko atim.  
I-scare-inv dog  
‘The dog scares me.’

However, we will assume that it is not so much the formal expression of these arguments themselves that plays a direct role in the choice of Subject, but their topicality. This factor then leads to pronominal rather than nominal expression of the referent. In other words, it is not the form which triggers the Subject rule. Form and Subject assignment are both determined by deeper pragmatic (and semantic) factors.

The fourth and final factor instrumental in assigning Subjects is mainly found in languages with a split alignment system, more specifically Accusative-Ergative. The choice for either of these systems is typically based on semantic aspects of the relevant arguments, mainly the prototypical Agent and Patient positions. However, in some of these languages other factors play a role in this decision. This is illustrated by an example from Georgian, in (18) below.

---


6. This is fully in compliance with our view on the expression of arguments in a functional theoretical framework. Cf. Siewierska and Bakker (2005).
Another take on the notion Subject

(18) **Georgian** (Kartvelian; Comrie 1978)
    a. Student-i ceril-s cers.
       student-nom letter-acc write:pres
       ‘The student writes the letter.’
    b. Student-ma ceril-i dacera.
       student-erg letter-abs write:past
       ‘The student wrote the letter.’

In this language, the choice for either of the alignment subsystems is based on Tense, among other factors.7 Thus in (18a), with present tense, the alignment is Accusative, while in the past tense, illustrated in (18b), the alignment is Ergative. Indirectly, this has a bearing on the choice of Subject. While the relevant criteria for Subject assignment in these examples – case marking and constituent order – conspire in (18a) to render the Agent Subject, they are in competition in (18b), leaving us the choice of either abandoning the notion, or developing a criterion to choose. We will opt for the latter, as discussed in Section 4 below. We will call a factor like Tense in Georgian contextual, since it does not directly emanate from the referents in the relevant argument positions.

This general sketch of the factors determining Subject assignment may be summarized as in Figure 1 below. On the lefthand side of this figure we find the languages for which all MSP are controlled by individual semantic functions. Acehnese was a potential example of such a language. On the righthand side we find the languages where pragmatic functions control the relevant phenomena. Tibeto-Birman languages like Mongsen may provide an example of these. In the middle we find the languages – arguably the vast majority - where the MSP are controlled by a set of factors drawn from the four domains mentioned, i.e. semantic and pragmatic functions plus the semantic features of the arguments and contextual aspects.

Figure 1 just represents the domains from which the factors for Subject assignment derive in a static fashion. In practice, these factors should be seen both as the motivating principles for Subject choice and the constraints on it. Within a language, competition between all these principles leads to a specific subset of well-formed constellations, i.e. an instantiation of Figure 1 on the basis of concrete values for the respective parameters. Constant competition may lead to diachronic changes in these instantiations.

In the next section we will have a brief look at some more complex cases of Subject assignment, allowing us to fill in some details of the very general picture sketched above.

---

7. See Dixon (1994:104f.) for a number of complex conditions on Ergativity.
4. Some complex cases of Subject assignment

So far it has been tacitly assumed that the relevant morpho-syntactic phenomena are relevant uniformly for all Subjects in a language, including their feeding into rules such as passivization. However, this is not necessarily the case. There exist a large number of different splits, incongruencies and restrictions in the area of Subject assignment within and across languages. Given the scope of this article, we will not make any attempt at completeness; rather, we will discuss a few of the more well-known cases, which we think serve to support our general view.

A first example are the so-called split ergatives. As already shown in the Georgian example in (18) above, languages may have several parallel ways of organizing grammatical relations. Not surprisingly, the factors that determine the choice between these systems stem from the same four domains which were mentioned as instrumental for Subject assignment in the first place. Thus in Georgian, with Accusative alignment as in (18a), both case marking and constituent order seem to indicate the Agent as the Subject. However, in (18b) case marking would favour the Patient whereas constituent order would indicate the Agent. Rather than abandoning the Subject notion in such cases, we propose a hierarchy of factors for establishing the best candidate. It looks as follows:

(19) MSP hierarchy (first version)

Morphological coding > Constituent Order

We thus suggest that in case of conflict case marking and agreement marking are better candidates for establishing Subjecthood than constituent order at the clause level. As shown in Siewierska et al. (1997), as well as in many other places, clause-level constituent order is without any doubt the most flexible of the traditional order parameters, and seems therefore to be the least reliably and diachronically stable of the respective domains of Subject diagnostics. Also, by contrast with morphological marking, constituent order is not a device that may or may not be developed by languages.
Another take on the notion Subject

It is trivially always there, and may or may not be employed for certain communicative goals.

As a second case, we will have a brief look at inverse languages, such as Nocte, as introduced above. Taking examples (16a/b) as our only data, and looking at the relevant morphosyntactic aspects, we have both the arguments for Subject candidates in (16a). The Agent has verb agreement and is fronted, while the Patient is unmarked. The same holds for (16b). However, the Agent is in first position, but agreement and zero marking now go to the Patient. Disregarding syntactic control structures for the sake of argument, we would consider the Agent for Subject in (16a) and the Patient in (16b). Not only do these constituents have two out of three features in their respective sentences, they also control the strongest of the morphological features, i.e. verb agreement, which has a syntactic long-distance aspect which is absent in zero marking. Thus, Nocte follows the same pattern as Georgian.

A third case, extensively studied in Van Valin and LaPolla (1997:356f.), is Icelandic. Consider the following examples:

(20) **Icelandic** (Indo-European; Zaenen et al. 1985)

a. Ég hjálpa-ð-i þeim.
   1SG.NOM help-PAST-1SG 3PL.DAT
   ‘I helped them.’

b. Þeim va-r hjálp-að af þér.
   3PL.DAT be-PAST-3SG help-PASTPRT by 1SG.DAT
   ‘They were helped by me.’

c. Þeim líka-r matur-in-n og bóð-a mikið.
   3PL.DAT like-3SG.PRES food-DEF-SG.NOM and eat-3PL.PRES much
   ‘They like food and eat much.’

Icelandic has the usual Subject features, as shown in (20a). Subjects take the (unmarked) Nominative case, and their Person and Number are marked on the verb. Icelandic has a regular passive, which promotes the Patient-like argument to Subject position, and demotes the Agent, to the extent that it is present in the first place, to an Oblique. In both cases, the Subject controls conjunction reduction. However, in the case of transitives with a dative second argument, which should probably be interpreted as a Recipient rather than a Patient, passivization has a rather different set of side effects, as shown in (20b). Indeed, in such cases the Dative argument is fronted, and the Agent gets Oblique status. However, the potential new Subject does not neutralize its case to Nominative, and does not copy its Person and Number features onto the verb. The latter takes 3rd Person marking, the form typically found in so-called impersonal passives across languages. Interestingly, however, conjunction reduction is controlled by this dative argument, as shown in (20c). From our perspective, this is enough to consider the Dative argument a Subject, be it a weaker version than the one for a certain purpose.
in the Active case.\textsuperscript{8} The marking of the verb is taken care of by a default rule. Dative marking takes place by the same rule that takes care of it in active contexts. It is simply not neutralized, as in the case of Accusative arguments being assigned Subject.

We think there are good reasons to consider “long-distance” syntactic control such as conjunction reduction and equi deletion to be more solid proof of Subjecthood than more local coding phenomena. In the light of this, then, we extend our hierarchy of (19) above to the form in (19’):

(19’) MSP hierarchy (extended)

\begin{itemize}
\item Syntactic Control > Morphological coding > Constituent Order
\end{itemize}

We could then say that in Icelandic, Subjects do not neutralize the case marking of non-Patients, nor do they mark them on the verb. However, they will be fronted and also have control over conjunction reduction. In other words, they feed only partially into passivization, and will therefore still be seen as Subjects, be it weaker ones.

Things are more complex than this, however. Consider the following examples from Dyirbal.

(21) \textit{Dyirbal} (Australian; Dixon 1994:12f.)

\begin{itemize}
\item a. N\textsuperscript{u}rra ngana-na bura-n.
  2\textsc{pl.nom} 1\textsc{pl.acc} \textsc{see-nonfut}
  ‘You saw us.’
\item b. Nguma yabu-nggu bura-n.
  father.\textsc{abs} mother-\textsc{erg} see-\textsc{nonfut}
  ‘Father saw mother.’
\item c. N\textsuperscript{u}rra ngana-na bura-n banaga-n\textsuperscript{u}.
  2\textsc{pl.nom} 1\textsc{pl.acc} \textsc{see-nonfut return-nonfut}
  ‘You saw us and we returned.’
\item d. Ngana bural-ngga-n\textsuperscript{u} n\textsuperscript{u}rra-ngu banaga-n\textsuperscript{u}.
  1\textsc{pl.nom} see-\textsc{antipass-nonfut} 2\textsc{sg.dat} return-\textsc{nonfut}
  ‘We saw you and we returned.’
\end{itemize}

If we consider only (21a) and (21b), Dyirbal appears to manifest itself as a split ergative language, with the (pro)nominal state of the relevant arguments as the determining factor for choosing between the subsystems, much in the way of Cree in (17) above. However, in the light of more complex constructions, we may have to reconsider this view. Although there may be no doubt about Subjecthood in (21a), in the case of conjoined clauses as in (21c), conjunction reduction turns out to be controlled by the Patient rather than the Agent. Under application of the Antipassive, control over conjunction reduction would go to the Agent, as in (21d). Note that this is effected by the demotion of the default Subject, the Patient, to a postverbal position, where it is marked by the Dative rather than the Accusative. We could say that default Subject

\textsuperscript{8} Such examples from Icelandic were already discussed by Cole et al. (1980).
Another take on the notion Subject assignment in Dyirbal is to the Patient in all cases. In the case of two non-topical – and thus nominally expressed – arguments the case marking follows the Ergative alignment, with an unmarked Patient which has control over conjunction reduction. The anti-passive shifts both to the Agent, which is then the Subject. In the case of two topical – and therefore pronominal – arguments the default assignment of Subject is again to the Patient. However, case marking follows a semantic pattern, more or less like the case of Icelandic dative predicates in (20). The effect of the Antipassive leaves the marking of the Agent unaffected, but it now assumes the Subject function. The Patient gets the oblique Dative case. Thus, we have Subject effective at the highest of the three domains, and case marking being determined partially by semantic functions, i.e. in the case of topical arguments. The order of the arguments may be fully determined on the basis of the semantic functions. The Agent is always in the first position, independent of its Subjecthood, and the Patient is preverbal when Subject and postverbal when not.

A final example we would like to discuss is from another Australian language, Yidin'. According to Dixon (1977), this language has ergative alignment for nominal NPs, and accusative alignment for pronominal ones. However, other than the case of Dyirbal in (21), both the Agent and the Patient may control conjunction reduction. This is exemplified in (22) below.

(22) Yidin' (Dixon 1977:390)
Ngayu bama bandar wandan.
1sg.nom person.acc follow fall
'I followed the person and I/he fell.'

This utterance unites the unmarked arguments from the two alignment systems, i.e. both Subject candidates. Indeed, each of them may be interpreted as the controller of the reduced Subject position in the second conjunct. Obviously, there will hardly ever be real ambiguity for hearers in these cases: the most topical of the two referents will be the natural controller. In a linguistic model restricted to the sentence level, such cases remain unsolved.

We are convinced that this does not in any way exhaust the possible constellations relevant for Subject assignment in the languages of the world. However, our framework might offer a starting point for rethinking the treatment of Subjects in a functionally oriented theory.

5. Conclusions

In this contribution, we have tried to sketch a general framework for the assignment of Subjects within languages. We started out with spoken data from the British National
Corpus. We showed that, in the case of passives with agents, the choice of Subject was never a ‘free’ choice of the speaker. In all attested cases the choice seemed to be based on pragmatic, and to a lesser extent, semantic factors of the relevant arguments. Extending our observations to other languages, we summarized the four domains from which these factors stem. However, a static, synchronic picture is an abstraction which may describe but not explain what Subjects in fact are. Thus, we think that choice of Subject is part of the dynamic, performance aspect of a language, based on a complex set of pragmatic and semantic factors rather than a more or less compulsory but speaker-controlled choice from among two or more candidates. Interestingly, since the rules for the selection of Subjects seem to operate mainly stochastically rather than in a yes/no fashion, they work in two ways vis à vis competence. On the one hand, they create constraints on what would logically be possible within the limits set by the grammar in terms of the Semantic Function Hierarchy and other devices. In that sense they work as a filtering device, be it a functional rather than a formal filter. Competition between the respective forces, together with other properties of the respective grammars, leads to different outcomes for different languages. Native speaker observations concerning the limits to Subject assignment in some language which are based on individual sentences-out-of-context are typically interpreted as what competence has to say about Subject. On the other hand, these forces in competition tempt speakers to cross the boundaries of competence, and to produce utterances which strictly speaking would be considered un-well-formed. Examples from Dutch of such transgressions are found in (23) and (24) below.

(23) De reiziger-s word-en verzocht uit te stappen.
    def passenger-pl aux-pass-pl request-pastprt get.off
    ‘Passengers are requested to get off the train.’

(24) Die broek pas ik niet!
    dem pair.of.pants fit.1sg 1sg not
    ‘Those pants do not fit me.’

In (23), Subject is assigned to the Recipient, which is normally outside the bounds of the Dutch SFH. However, (23) is frequently heard spoken by conductors on trains. Given the topicality of passengers in such a context, and the contrast between the concreteness of the referent of the Recipient and the abstractness, and total unfitness for Subjecthood, of the Patient argument, this form has become a fully acceptable utterance in such contexts. It is corrected only by schoolteachers and the like. In (24), the same features which are found to be active in Subject assignment, such as Person and Animacy, seem to be in the process of bringing about a reinterpretation of the predicate scheme of the verb passen ‘to fit’. In its standard dictionary reading, the first argument would be the Patient (here: the pair of pants) and the second one the Experiencer (here: the first person). However, younger speakers of Dutch will typically reverse the roles, and interpret the Experiencer as the first argument, and therefore make it the default
Subject, while the original meaning is maintained. Without making performance factors central, there would be no way apart from acquisition errors to explain these and many other diachronic changes. And without taking performance, and discourse, into consideration it is probably impossible to explain and describe what speakers do when linguists say that they assign Subject.

Abbreviations

| 1 | first person | LOC | locative |
| 2 | second person | MASC | masculine |
| 3 | third person | NEG | negative |
|-abs | absolutive | NOM | nominative |
|-acc | accusative | NONFUT | nonfuture |
|-act | active | NRP | nominalizer prefix |
|-ag | agent | PASS | passive |
|-antipass | antipassive | PAST | past tense |
|-aux | auxiliary | PASTPRT | past participle |
|-ben | beneficiary | PAT | patient |
|-def | definite | POSS | possessive |
|-dem | demonstrative | PRES | present |
|-dir | direct | PROX | proximate demonstrative |
|-erg | ergative | PRT | particle |
|-fem | feminine | REC | recipient |
|-instr | instrument | SEQ | sequential converb |
|-inv | inverse | SG | singular |

References


10. Older speakers would interpret (24) as a refusal to try and put on the pants, since there is a second meaning of the same verb i.e. ‘see whether it fits’ with a more prototypical Agent-Patient argument structure.


The modal auxiliaries of English, $\pi$-operators in Functional Grammar and “grounding”*

Louis Goossens
University of Antwerp

In this contribution it is argued that in spite of their homogeneous morphological and syntactic properties, there are compelling reasons to differentiate the central modal auxiliaries of English with respect to grammaticalization. The main differential semantic criterion comes from their contribution as grounding predications, as understood by Cognitive Grammar. The implications of this differentiation will mainly be investigated for their treatment (as $\pi$-operators or otherwise) in Functional Grammar, but it will also entail a modified view for their status as grounding predications in Cognitive Grammar (as compared with Langacker’s). Insights from the diachronic development of must are taken to support the position defended here.

1. Introduction

This paper is concerned with the grammaticalization status of the modal auxiliaries of English, as well as with the ways in which this status is to be understood within Standard Functional Grammar (Dik 1997) and Cognitive Grammar (Langacker 1991).

It will be argued that the degree of grammaticalization of the central English modal auxiliaries is not the same in all their uses. Although they form a fairly homogeneous set from a formal and syntactic point of view, they exhibit semantic differences which are too important to be ignored in a discussion of their status as grammaticalized items.

With respect to the way in which those differences have to be accounted for within the model of Functional Grammar (henceforth FG) it follows in my opinion that the standard representation in terms of different $\pi$-operators, i.e. operators on the nuclear

---

* This paper is an elaborated and considerably revised version of a contribution written in French which was presented at the colloquy “Les développements récents de la Grammaire Fonctionelle” (Mohammedia, 13–14 April 1999), and which appeared as Goossens (2003).
predication with different scopes, will have to be reviewed critically: only fully grammaticalized uses deserve operator status. Obviously, this requires a semantic criterion which can be used to draw the line between those uses that can be taken to be fully grammaticalized, and those that are not.

In our quest for such a criterion a concept from Cognitive Grammar (CG) is invoked. ‘Grounding’ is a key-notion in CG which is used to account for the ways in which tense and modality function in clauses. As will become clear, however, my interpretation of this function is more consistently semantic than in Langacker’s (1991) application to the English modals. As such, it will help to clarify our understanding of the central English modals, as well as of the question to what extent those central English modals should be given a representation by means of π-operators or not.

Successively, we present a discussion of criteria relating to grammaticalization, especially with respect to the English modals (Section 2), of π-operators in FG and of standard FG’s approach to modality and the English modals (Section 3), and of ‘grounding’ as well as of the relevance of this concept for the modal auxiliaries as conceived by Langacker (Section 4). A modified view about grounding and a graded grammaticalization account of the English modals come in Section 5. By way of conclusion we round off with the implications of all this for a treatment of the modals in FG as well as in CG (Section 6).

2. Criteria for grammaticalization and the modal auxiliaries of English

2.1 Criteria for grammaticalization

Basically there are two sorts of criteria which play a part in the decision whether a given linguistic expression is grammaticalized. On the one hand there are formal criteria. Bound morphemes, e.g. the morpheme for the past tense in English, can always be taken to be grammaticalized. Free morphemes which belong to a closed class with special syntactic characteristics are usually good candidates. The modal auxiliaries of Present-day English may be taken to belong here. Specific word order templates (such as inversion), often in combination with certain intonation patterns, may also be considered as instances of grammaticalization.

This is not the whole story, however. The semantic correlate of formally grammaticalized items is typically a meaning which is quite abstract (or schematic, to use another terminology), and as a rule either quantificational (e.g. singular-plural opposition, perfective:imperfective, semelfactive:repetitive) or tied up with what Langacker calls the ground and as such subjectified (e.g. Tense oppositions, the expression of ‘basic’ speech acts; a fuller discussion of grounding follows in Section 4).

As a rule the formal and the semantic criteria go hand in hand; this is certainly the case for bound morphemes. But in the case of free morphemes belonging to a closed class the meaning criterion may come in to determine degrees of grammaticalization. One instance that may be invoked here is the behaviour of have in British English.
Although in its possession sense have may (still) go without do-support in a restricted number of contexts such as those in (1) and (2), one would not want to put those instances on a par with perfect have as regards grammaticalization, not just for formal, but also for semantic reasons: they do not have ‘grounding’ meanings as opposed to perfect have.

(1) I haven’t any money.

(2) He hasn’t a very good temper.

Note also that the notion closed versus open class is graded. Temporal adverbials with a deictic character, for example, are considerably more restricted than non-deictic temporal adverbials, but the fact that we still very much get an open-ended class (now, yesterday, tomorrow, last night, last week, next week, next month, the day after tomorrow, etc.) does not make them clear instances of grammaticalization, in spite of their ‘grounding’ contribution. For another thing their meaning is rather explicit, and therefore not ‘subjectified’ in Langacker’s terminology.

Again, we do not go into details here, but turn to our showcase, the central modals of English.

2.2 English modals and grammaticalization

The central modals of English are normally defined by the following formal characteristics:
– they are as a rule combined with the to-less infinitive of another verb
– the third person singular does not take -s
– they have no participles or infinitives
– they do not receive do-support in negative, interrogative, emphatic sentences or in short substitute sentences.

On the basis of these formal features the set of modal verbs/auxiliaries is taken to be composed of the following items: can—could, may—might, will—would, shall—should and must (ought is not a central member of this set, because it is followed by a to-infinitive, and permits an increasing degree of do-support in Present-day English).

On the other hand, the ‘central’ modals are considerably diversified with respect to the meanings which they express:
– some express facultative modality (capacity/ability: can, could; volition/willingness: will, would)
– secondly, there are the deontic uses (permission: can, could, may, and marginally, might; obligation: must, should, and, occasionally, shall)
– inferential or epistemic modality is expressed by may, might, could, should, will, and occasionally, would)
– the future is expressed by will and, with restrictions, shall
– conditional meaning may be signalled by should and would.
The question that may be asked here is whether this semantic diversification has implications for their status as grammaticalized items. One might indeed argue that some of these meanings are less abstract or schematic than others, and that this has consequences for their status as \( \pi \)-operators in FG. To see how this problem can be approached we first summarize the treatment of modality (and of the English modals) in Dik (1997) (Section 3). Next I proceed to a discussion of ‘grounding’ as understood by Langacker, and consider its relevance for the problem at hand (Section 4).

3. \( \Pi \)-operators and modality in Functional Grammar

3.1 \( \Pi \)-operators

FG operates with underlying structures, which are then turned into actual natural language sentences by the expression component. The underlying structure of a given sentence tries to give an account of all its meanings (including all its different functions). The starting-point of such a sentence is as a rule the so-called nuclear predication, which typically includes a predicate and its arguments.\(^1\) Such a nuclear predication can then be extended by \( \pi \)-operators of four different levels, each of them representing a grammatical extension (Dik 1997: Part 1, Chapters 9 and 12); each higher level has the lower level within its scope. In this sense FG is a ‘layered model’.

i. Extension of the nuclear predication by \( \pi \)-1 operators results in the "core predication" (an example of a \( \pi \)-1 operator is \textit{Prog}, which produces the Progressive in English).

ii. Extension by \( \pi \)-2 gives us the “extended predication” (Tense operators belong here, e.g. \textit{Pres} in English to produce the Present Tense).

iii. Extension by \( \pi \)-3 gives us the “proposition” (\textit{Poss} for Epistemic Possibility as typically expressed by \textit{may} in English).

iv. Finally we get \( \pi \)-4, which produces the “clause” (here we get the realization of the (basic) speech act, e.g. \textit{Decl}, for Declarative, which in English is expressed by the typical word order(s) of a statement).

The layered (underlying) structure of the clause and the associated presence of \( \pi \)-operators is an indispensable and fundamental notion in FG. But it is also clear that in its standard form it is not unproblematic.

i. One problem is the exact number of layers that is required. To quote just one contribution that raises this problem we refer to Moutaouakil (1996), who has argued that there are several reasons to believe that the level of the proposition

---

1. Predicates may be verbal or non-verbal. Moreover, FG also provides for the possibility of incomplete or holophrastic expressions. The details of all this, however, are not relevant to our purposes.
The modal auxiliaries of English, π-operators in Functional Grammar and “grounding” 163

should be limited to the declarative clause and does not function in interrogative, imperative, or exclamative clauses.

ii. The strict separation among the different layers as envisaged in standard FG poses another problem. One might argue, for example, as I have done in an unpublished paper presented at the “Sixth International Conference on Functional Grammar” (York, 1994), that the operator Fut (Future) could be assigned to both Level 2 and Level 3. As an operator of Tense it belongs to Level 2, but, since the expression of futurity may be interpreted as a judgement by the speaker concerning a state of affairs which has not yet been realized, one may also take it to be an epistemic qualification and assign it to Level 3.

iii. Another problem which is not envisaged by the standard theory is one which I have signalled in Goossens (1994). In that paper I have demonstrated that operators may have variable scope. In an instance like (3) the (quantitative) operator Rep(etitive) has scope over the (aspectual) operator Imp(erfective) (or Prog(ressive)), whereas in (4) the scope is the other way round.

(3) They were watching TV in the dining-room whenever I dropped in unexpectedly.

(4) In those days they were eating breakfast in the dining-room.

iv. The problem which concerns us in this paper relates to the question how the fact that grammaticalization is a graded rather than a black or white issue affects our view of layering. The illustration comes from the (central) English modals, although other instances might be adduced. Just to mention one which was brought to my attention by Lachlan Mackenzie when I was working on Goossens (1994), the behaviour of English used to as a marker of repetition or habit is only ‘half’ grammaticalized in Present-day English, given the developments in do-support in recent English. But, as stated, our focus here is on the modal auxiliaries of English, and as it will turn out, the argumentation has to do both with formal and semantic considerations.

3.2 Modals and modality in Dik (1997)

It is clear that Dik, as is shown by the following quotation, takes a diversified approach to modality and the modals: “we distinguish a number of modality types, which we assign to different parts of the structure of the clause and correspondingly, to different parts of the clause. We shall reserve the term Mood for those modality distinctions which are grammatically expressed (cf. Aspectuality vs. Aspect, Temporality vs. Tense)” (Dik 1997:241).

It is not clear, however, at what point a modal is considered to be grammatically expressed, or what is meant exactly by placing inherent modality (what I referred to as facultative modality above) and certain deontic uses at Level 1. I quote the passage in full.
Level 1. *Inherent modality* distinctions define relations between a participant and the realization of the SoA in which he is involved. These distinctions may consist in the ability or the willingness of a participant to do the SoA (*can, be able to/want, be willing to*) or permitted (*may, be allowed to*) to do the SoA. Inherent modalities are not expressed by grammatical means, and belong to the internal structure of the predication. They are mainly mentioned under Modality because the predicates used to express these SoA inherent features often develop into more strictly “modal” expressions over time. (Dik 1997:241–242)

Although I largely agree with what is stated here, I have two difficulties with it. First, ‘Level 1’ would seem to imply a treatment by means of π₁ operators (on page 242 ‘Level 2’ is correlated with π₂ operators, and ‘Level 3’ with π₃ operators). On the other hand, the statement that they belong to the internal structure of the predication implies that these predicates have an argument structure of their own, and are not expressed by an operator. Second, and more importantly with respect to the topic of this paper, the statement that they are not expressed by grammatical means is not obvious for the central modals which are explicitly mentioned here, viz. *can, may* and *must*, because at least from a formal point of view they are not distinguishable from their other uses or from the other central modals (whatever their meaning may be). It would seem to me that in this discussion the decision about their grammatical status is entirely made dependent on their meaning, regardless of their formal behaviour.

We now turn to Cognitive Grammar where a different line of thinking is followed.

4. Grounding

4.1 Grounding: A basic notion in Cognitive Grammar

Grounding is “[a] semantic function that constitutes the final step in the formation of a nominal or a finite clause. With respect to fundamental “epistemic” notions (e.g. definiteness for nominals, tense/modality for clauses), it establishes the location vis-à-vis the ground of the thing or process serving as the nominal or clausal profile” (Langacker 1991:549). The ground is “[t]he speech event, its participants, and its immediate circumstances” (Langacker 1991:548).

For Langacker, an essential property of grounding predications is that G is “necessarily implicit despite its pivotal role as a reference point” (Langacker 1990:15). Consequently, a phrase like “I think”, which undoubtedly expresses the opinion of the speaker (a central ingredient of the ground) is not “grounding” because the reference to the speaker is explicit. A sentence like “John may have done it”, on the other hand, expresses speaker opinion in an implicit way and is grounding.

A correlate of this property of implicitness is that “grounding predications” are to be taken to be expressed by grammaticalized items, which (as a defining feature) have schematic meanings. As Langacker notes, “we can reasonably suppose that only “grammaticized” (as opposed to “lexical”) elements can serve as true grounding pred-
ications. For the moment, it is sufficient to note that a grammaticized element is quite schematic semantically (i.e. it lacks the specificity and rich detail typical of lexical items) [...]” (Langacker 1990:13).

From a diachronic point of view, but also in an interpretation of synchrony as essentially dynamic, another important notion is that of ‘subjectification’. Subjectification “is a semantic shift or extension in which an entity originally construed objectively comes to receive a more subjective construal” (Langacker 1991:215). This involves an implicit reference to the speaker (or to the ground). The process of subjectification can be illustrated by the history of the central modals, which have developed from full lexical verbs to the modal auxiliaries of Present-day English for which subjectification as understood by Langacker is a clear feature.

Note that the correlation grounding-grammaticalization-subjectification is not defended by all linguists who are concerned with subjectification. For example, in Traugott and Dasher (2002), where the role of subjectification in semantic change is the central issue, subjectification and grammaticalization do not have to go hand in hand. Still, for them, too, the modals (in English as well as in other languages) provide a crown example to illustrate the process of subjectification. In this paper, however, I go along with Cognitive Grammar, but not without certain qualifications of Langacker’s position.

4.2 Grounding and the English modals in Cognitive Grammar

For Langacker, the English modals are “grounding predications” (1991:271). His comparison with the German modals shows that this characterization is largely based on formal characteristics: “That the German modals are not grounding predications is shown by several differences between their grammatical behavior and that of their English counterparts. For one thing, they all have infinitival forms, which is inconsistent with the function of grounding a finite clause [...] They also form past participles that occur in the perfect construction as complements to the auxiliary verb ‘have’ [...] Furthermore, the German modals agree in person with the subject [...]” (1991:271).

At the same time he recognizes that there is a gradation of subjectivity when he notes that “we have outlined a developmental sequence – leading from main verbs, through root modals of various sorts, to epistemic modals – whose organizing property is that the locus of potency is construed with progressively greater subjectivity at each step along the path” (Langacker 1991:275). On the other hand, he does maintain the position that the central modals are grounding predications in all their uses. In what follows it will be argued that there are (semantic) arguments to consider some of their uses as less grammaticalized, and not subjectified (or ‘grounding’), in spite of the fact that their formal behaviour is identical to that of the subjectified ones.
5. Grounding as an argument for a graded view of the grammaticalization status of the English modals

5.1 Grounding and tense

According to Langacker (1991: Chapter 6) grounding in the verb phrase is effected by tense or by modality. Tense and modality are complementary, although in the expression of the future, which in English is expressed by means of modals (mostly WILL, occasionally by SHALL) there is some interpenetration of modality and tense. Tense operates in Reality, either Immediate Reality or (Known) Reality preceding Immediate Reality (see Fig. 6.1 in Langacker 1991:242). If the grounding operation is effected by modality it is not effected by tense. On the other hand, if it is effected by tense and if in addition there is an expression of modality which combines with it, then tense has the modality in its scope.

Both tense and modality can be regarded as grounding operations. Tense places the predication over which it has scope in known reality (as a rule, though not always, by marking it as either simultaneous with the moment of speaking, i.e. as Present, or as preceding the moment of speaking, i.e. as Past). Modality places it in ‘irreality’; modality creates some sort of epistemic distance. For Langacker the central modals in Present-day English have this grounding effect without exceptions. In what follows I would like to defend the position that they do in the majority of their uses, but not in all of them (see also Goossens 1996).

5.2 Subjectified, grounding uses of the central modals

Normally the central modals of English are subjectified, and express a meaning that reflects the ground. These uses cannot be tensed (they are not marked for Present or Past).

i. Instances like (5) express subjective, epistemic modality. They are ‘grounded’ by this modality and not understood as tensed. In FG they would be represented by a $\pi$-operator of Level 3.

(5) He may/might/must/could/will have seen her.

ii. In instance (6) we have the expression of a deontic modality which implies speaker authority (the speaker is understood to be the authority source). One may argue that the modal has a direct directive force. In FG this would have to be represented by means of an operator at the level of the illocution ($\pi$-4). At any rate, it is clear that these uses cannot be tensed.

---

2. A similar argumentation is to be found for the majority of uses of du musst in Present-day German in Mortelmans (1999: Section VII.1).
(6) You must/may do this exercise twice a day.

iii. As I have already observed before, the use of *will* (and to some extent of *shall* in the first person) to form the so-called Future Tense, as in (7), has a grounding function in which tense and modality interpenetrate.

(7) They will travel together.

*Will* puts the predication within a predictable, and to that extent known reality, it grounds the predication in the time span after the moment of speaking. At the same time, however, the very contingency of every prediction creates epistemic distance, which is why Langacker places it in Irreality. Obviously this is just another type of grounding. Technically in FG this may call for either a π-2 operator or a π-3 operator, or both. What I want to emphasize here is that tensing can be regarded as implying an epistemic qualification which signals that the predication belongs to known reality, and that Tense and Epistemic modality reflect two sides of the epistemic coin.

iv. Finally, the typical value associated with *would* in (8) (and with certain uses of *should* in the first person) is the expression of the conditional, which effects grounding in a mental space which is marked by the speaker as dependent on an unfulfilled condition (cf. Fauconnier 1994) and as such irreal.

(8) She would be happy if you did it.

In FG this could be represented by a π-operator Cond(itional) to be assigned to Level 3.

5.3 Non-subjectified, non-grounding uses of the central modals

The following uses of the central modals, on the other hand, do permit tensing. More particularly, the opposition Present:Past remains operational, and the grounding is effected by placing the predication in immediate reality (the time of speaking) or in the (known) reality preceding the time of speaking (there is in other words no epistemic distancing).

i. *Can:could* expressing facultative modality (‘subject-internal modality’ in the terminology employed by van der Auwera & Plungian 1998) as in (9) and (10).

(9) John can jump across this ditch (his capacity is present and real; the fact that an actual jumping act is only potential makes this use ‘modal’, but that does not invalidate that the grounding is one that places this capacity in immediate reality).

(10) John could jump across this ditch when he was your age (here the capacity is placed in the (known) reality preceding immediate reality, even if John may never have jumped across in the past; if we want to emphasize that John actu-
alized his jumping capacities in a jumping act, a phrase like ‘was able’ would have to be used).

ii. Although this use is somewhat restricted, WILL:WOULD can still be used in its original volitional sense; the Present:Past opposition is possible, especially in negative uses such as in (11) and (12).

(11) John will not (won’t) help, he is in a bad mood (this use is ambiguous between a predictive and a volitional interpretation, but the volitional interpretation is the one intended here).

(12) John wouldn’t help, he was in a bad mood (refusal in the Past, i.e. known reality).

iii. Another use of can:could which is related to that under (i), but not identical with it, because it expresses what van der Auwera and Plungian refer to as ‘subject-external modality’, is exemplified by (13) and (14).

(13) We can make coffee upstairs (to some extent this involves a Present ability, but the enablement is very much tied up with external circumstances).

(14) We could make coffee upstairs (a possibility in the past, otherwise similar to that in (13)).

iv. Fully subject-external is the use of can:could to express permission. Again the opposition between (15) and (16) is one between past and present.

(15) We can keep our shoes on (present permission).

(16) We could keep our shoes on (past permission).

In all of these uses, (i)–(iv), grounding is effected through tense, not through modality, and the modals concerned are not subjectified. Somehow, however, the special morpho-syntactic properties which these uses share with the grounding uses must be taken care of. In FG the most obvious possibility to account for these non-grounding uses would be to deal with them, not as π-operators, but as involving predicate formation rules, as proposed in Goossens 1987 and worked out in more detail in Goossens 1992 (Section 4). In the latter paper I also proposed, be it somewhat tentatively, to deal with the ability use mentioned here under (i) as involving an independent predicate can with an argument structure of its own, but to be assigned to the category mv. (modal verb) rather than v. (verb) to do justice to its special morpho-syntactic characteristics.

5.4 Transitional uses (interpenetration of tense and modality)

It is to be expected that there are uses of the central modals for which it is difficult to decide whether they are grounding or not. The following uses with must illustrate this.
(17) Clay pots [...] must have some protection from severe weather (example borrowed from Coates 1983:35).

In this instance *must* is used deontically, but it is not clear whether speaker authority is involved. *Must* expresses some sort of general necessity; one might say that this necessity comes within the scope of a generalized type of Present, but a Present:Past opposition is ruled out for *must*. To mark it as Past one would have to resort to HAD TO.

The next example, which comes from the spoken LondonLund corpus, shows *must* in an inferential use, but one which is objective rather than subjective. Again, the Present:Past opposition would not be operative, but it is not a full-scale grounding use, in the sense that it does not fully express subjective speaker judgement.

(18) what Fan didn’t realize not being a lawyer # or a lawyer’s wife # – that apparently the wife was put on probation # so that Fan didn’t realize that # she must also have been up before court # you can’t be put on probation # not unless you’re guilty.

Although *may* is as a rule subjectified, there are similar uses of *may* which are rather objective, but where *might* could not function as a marker of the Past. The transitional nature of these usages makes it difficult to decide whether to deal with them as \(\pi\)-operators in FG, or perhaps as involving a predicate formation rule. This should not really be a difficulty for the theory of FG, if at least it is prepared to recognize these transitional uses for what they are.

Transitional cases, for that matter, pervade the history of the development of the modals: the diachronic development of *must* in the next subsection may serve as an example to illustrate the shift from non-subjectified to subjectified item, as well as the correlation with the loss of the Present:Past opposition.

5.5 The shift from non-grounding to grounding: The diachronic development of *must*

In Goossens (2000) I studied the development of *must* from a non-subjectified to a subjectified item on the basis of the (diachronic) Helsinki corpus. It appeared that this subjectification manifests itself first in the deontic domain. I found that in the last two samples for Middle English, ME3 (1350–1420) and ME4 (1420–1500), and in the first Modern English sample, EMOE1 (1500–1570), there is a constant rise in deontic instances in which the speaker’s authority is involved. But the majority of those deontic uses remain non-subjectified, in the sense that the speaker is not the source of the obligation or that (s)he does not associate him/herself with this authority source, in other words the majority of those deontic uses are non-grounding. It is in EMOE2 (1570–1640) that those subjectified uses constitute the majority for the first time. Furthermore, this subjectification of deontic *must* appears to have consequences for the epistemic domain. It is indeed the case that in EMOE3 (1640–1710) the objective-inferential uses also begin to be outnumbered by subjectified epistemic uses.
An important correlation from the point of view of the foregoing discussion in which the complementarity of grounding through tense and modal grounding was emphasized, is the fact that past tense uses of must disappear at the time when the subjectified modal uses (whether deontic or epistemic) begin to predominate. The last uses of past must (exemplified by (19)) occur in EMOE1 (1570–1640).

(19) [...] and departed agayne into Galile. And it was so that he must nedes goo thorowe Samaria. Then he came [...] (must would have to be rendered by had to in Present-day English).

At any rate, this diachronic excursion illustrates the gradual character of the change to grounding predications for must. Given the facts observed is 5.3 it should be clear that can has not progressed in the same way as a grounding predication, and that the central modals should not be taken to contribute all to the same degree to the grounding of verb phrase.

6. Implications

6.1 Implications for Cognitive Grammar

Looking at the central modals of English with respect to grounding as understood by Langacker, I found that they are indeed grounding predications in most of their uses. Taking into account, however, that grounding in the verb phrase is effected either through tense (the predication is placed in known reality as Present or Past) or through (subjectified) modality, I do not take the central modals to be grounding predications in all of their uses. As long as the grounding is taken care of by means of tense (as in the instances discussed in Section 5.3) it is tense and not the modality expressed (e.g. ability, permission or volition) which establishes the link with the ground.

This position differs from Langacker’s, who invokes the shared formal characteristics to accept an enlarged concept of grounding (‘somehow associated with the ground’). It seems to me, however, that those shared characteristics should not lead to an overstatement of their grounding impact. Moreover, if we look at the morphosyntactic features one by one, we see that they are largely the consequence of earlier developments. The absence of an s-morpheme in the third person of the singular follows from the fact they originated as preterite-presents in the case of can and may, as an older optative in the case of will, or as a past indicative or subjunctive in the case of the ‘oblique’ modals could, might, would and should. Must is also a preterite-present in origin, but has a more complex history, which I do not go into here. The fact that they have no non-finite forms also has antecedents in their history: they were already partially defective in Old English. Still, the complete loss was no doubt promoted as well by their increasing use as grounding predications (in the more restricted sense as I understand it). Indeed, if they are used to express the subjectified uses discussed in 5.2, they necessarily occur in their finite forms. In my view then the absence
of the non-finite forms for the uses of 5.3 is an instance of (formal) paradigm pressure and should not be explained by overstretching the concept of grounding. Finally, the absence of do-support can be accounted for along the lines of what I elaborated in Goossens (1984):3 “The fact that the modals can, may and will (which at the time still had their infinitives) were excluded from do-periphrasis in its initial stages, can be accounted for on semantic grounds because periphrastic do derives from a pattern where it combined with action verbs only. In its later stages do-periphrasis appears to have remained resistant for a long time against a combination with the class of verbs to which the modals belong” (1984:157).

Summing up, I am in basic agreement that the central modals are developing into grounding predications and that in the majority of their uses they have done so, but I do not agree that this evolution is fully complete, in spite of the fact that all the members of the group of central modals share the same morpho-syntactic characteristics.

6.2 Implications for Functional Grammar

In FG the formal characteristics of the central modals can be taken care of by grouping them together as a formal category of their own, viz. mv. (modal verb), to be distinguished from v. (verb). The initial question, however, was how they have to be accounted for in the layered structure of the clause. If we take it that grounding as understood in this paper is an important semantic correlate of grammaticalization, there is no doubt that the uses under 5.2 require the presence of π-operators in the underlying structure. Given the complementarity between Tense and Subjective Epistemic uses we might propose a fusion of π-2 and π-3, since they present us with two sides of the epistemic coin (known reality which is either Present or Past, and ‘irreality’ where subjective assessments of reality by the speaker are at stake). In this fused view of Levels 2 and 3, the problem of the Future Tense which is inherently an assessment of the speaker of reality-to-come is taken care of.

The uses mentioned in 5.3, in which the modal has no grounding function, should be dealt with differently. The quotation from Dik (1997) in Section 3.2 already pointed that way: Dik even appears to consider some of these uses as predicates in their own right, i.e. with an argument structure of their own. As I have pointed out, however, the formal features may be a reason to go for an underlying structure which brings out their partial grammaticalization and to represent them by means of a predicate formation rule (cf. my proposals in Goossens 1987, 1992 and 1996). The uses under 5.4 present as it were points of entry into a status as π-operators.

3. Note that the ancestors of must and shall were already defective when do-support was in the making.
6.3 Envoi

A final perspective that may have emerged from this discussion is that FG and CG can both profit from confrontatations like the foregoing. A crucial requirement, however, is the necessity to analyse the natural language data in sufficient depth. In the present case it also proved valuable (once more, as far as I am concerned) to interpret those data as part of an essentially dynamic system.

References


The modal auxiliaries of English, π-operators in Functional Grammar and “grounding” 173


The king is on huntunge

On the relation between progressive and absentive in Old and Early Modern English*

Casper de Groot
University of Amsterdam

This paper addresses the diachronic development of two periphrastic constructions in Old and Middle English, *He wæs huntende* and *He wæs on huntunge*, into the progressive in Modern English. The literature on the origin of the progressive offers several hypotheses for explaining the coalescence of the two constructions. This paper offers a new hypothesis based on the consideration that the first construction, consisting of *be* + present participle, developed into the progressive, and that the second construction, consisting of *be* + *on* + verbal noun, was originally a construction denoting absence. The evidence for the coalescence comes from a partial overlap in the semantics of the progressive and the absentive, and the fact that progressives often originate from spatial constructions.

1. The scope of the paper

It is a generally accepted view concerning the rise of the progressive in English that the form originates from constructions with present participles used as predicative adjuncts and constructions with *be* + verbal noun. Fischer (1992:251), for instance, states that the stages of development of the progressive in Middle English can still be seen in constructions where it is not clear whether the *-ing*-form is an adjective (1), an appositive participle (2), or whether it is a progressive.

1. What ladyes fairest been or best daunsynge, (*CT* I.2201 [I: 2203])
2. Heere is the queene of Fayerye, / With harpe and pipe and symphonye, / Dwellynge in this place. (*CT* VII.814-16 [10: 814-16])

*I would like to thank Olga Fischer for her attentive reading of my text and her insightful, critical comments.*
Fischer continues by saying that the reasons for the increase in the use of progressive forms are not clear. In the Early Middle English period their frequency is very low. At the end of the Middle English period the frequency almost doubles and from the beginning of the Modern English period onwards the use of the construction rises astronomically.

One possible explanation for the increase of the use of the progressive is based on the following observations made by Visser (1963–1973: par. 1018–1034). In the Late Old English / Early Middle English period the inflectional endings of the present participle, (inflectional) infinitive, and verbal noun began to be confused. The verbal noun (in Old English ending in -ung) began to develop verbal properties. It is suggested that in connection with a number of phonological changes (see Lass 1992:145) the two separate Old English constructions in (3) became very similar in Middle English, as in (4), and ultimately coalesced, thus sharply increasing the frequency of the progressive form proper.

\begin{align*}
(3) & \text{he was huntende and he was on huntunge} \\
(4) & \text{*he was huntyng(e) and he was on/an/in/a huntyng(e)}
\end{align*}

Fischer (1992:253), however, notes that it is difficult to ascertain such a chronological development in the actual examples. Nehls (1974:168) suggests the following scenario: the two types (i.e. verbal and nominal -ing as in (4)) were functionally almost equivalent; as a result mixed forms appeared, the progressive finally replaced the on hunting type, which became dialectal and/or non-standard. Dal (1952) postulates that there were two periphrases in Old English, one of a literary nature, the other of a more popular character:

\begin{quote}
In der altenglischen Zeit haben wir die zwei heterogenen Typen he was feohtende und he was on feohtinge. Die Konstruktion mit -ing/-ung-Nomen hat der niederen Sprache angehört, und ist nur sehr vereinzelt in die schriftliche Ueberlieferung eingedrungen . . . (quoted in Scheffer 1974: 246)
\end{quote}

This paper builds on the assumption that the constructions in (3) did indeed develop into the constructions in (4), which then merged into just one type, the present-day progressive. Given the development of two constructions into one, one may wonder what the original nature of the two different constructions was, in particular where they differed, and in what respect they were functionally almost equivalent. I will defend the view that the periphrastic construction with a preposition and the verbal noun was used to indicate somebody’s absence in the sense of “the king is off hunting” (absentive), whereas the other construction was used to express somebody’s involvement in an activity (progressive).

1. The idea of the fusion of two constructions in Middle English goes back at least to Jespersen (1949:168).
The data used for this purpose come from the existing linguistic literature. No efforts have been made to find other new illustrative examples from the original sources, i.e. the various Old and Middle English manuscripts.

2. Progressive and Absentive

In their paper on the Progressive in the languages of Europe, Bertinetto et al. (2000) distinguish between three types of constructions. I adopt their typology for the discussion of the English data, since it offers a perfect instrument to illustrate the relevant differences which hold between the constructions from different stages of English. The characterization of the three types is as follows (Bertinetto, Ebert, & De Groot 2000:527).

(5) a. Focalized progressive construction, i.e. those expressing the notion of an event viewed as going on at a single point in time; e.g. *When the lamp fell on the table, John was eating his porridge.*

b. Durative progressive construction, i.e. those that are evaluated relative to a larger interval of time; e.g. *Last Tuesday, Mary was giving talks in London.*

c. Absentive construction, i.e. those conveying the meaning of an event occurring in a place (characteristically reserved for a given purpose) displaced from the deictic centre; e.g. *Peter is off buying tickets for the concert.*

Present-day English does not make a threefold distinction, but rather a distinction between (5a) and (5b). The progressive form is employed in both the Focalized progressive and the Durative progressive constructions. An Absentive construction, however, does not occur. Constructions as in (5c), with an overt expression of absence, do not count as grammatical expressions of absence. As we will see below, the situation in Early Middle English seems to differ in that the language makes a distinction between a Durative progressive (5b) and an Absentive (5c) one.

Note that the major difference between the progressive constructions on the one hand and the absentive on the other is that the progressive relates to temporal deixis whereas the absentive relates to spatial deixis. The use of the progressive in English may have a spatial interpretation, but it cannot be considered to be a grammatical expression of absence such as we find in for instance Dutch, Fering, German, Norwegian, Swedish, Italian, Hungarian and Finnish. The answer to a question of the type “Where is John?”, where somebody inquires after a place and not an activity, may contain a

2. English does not have an Absentive in the strict sense, since the particle *off* is used.

progressive in English, but not in Dutch (nor in the other languages with a grammatical expression of absence). In Dutch and the other languages, an answer using the Absentive would be appropriate (6b), whereas the progressive would not (6a).

(6) **Dutch**

a. Progressive
   
   Jan is aan het boksen.
   
   John is at the box.inf
   
   ‘John is boxing.’

b. Absentive
   
   Jan is boksen.
   
   John is box.inf
   
   ‘John is off boxing.’

The following examples from Dutch show that from a deictic perspective, the progressive and absentive form a minimal pair. When using the progressive in the main clause (7a), there is a collocation of the action (‘boxing’) and the deictic centre (‘room’), whereas there is a dislocation if the absentive is used (7b). In the latter case, Peter will find the room empty, while John is away boxing:

(7) **Dutch**

a. Progressive
   
   Toen Peter de kamer binnenkwam, was Jan aan het boksen.
   
   when Peter the room came_in sat John at the box.inf
   
   ‘When Peter entered the room, John was boxing (there).’

b. Absentive
   
   Toen Peter de kamer binnenkwam, was Jan boksen.
   
   when Peter the room came_in was John box.inf
   
   ‘When Peter entered the room, John was off boxing.’

Before I return to the English data, a few words are needed about the Absentive. The Absentive, as attested in at least eight languages of Europe (cf. De Groot 2000), is a grammatical expression of absence. As the examples (6b) and (7b) show, there is no overt element in the expression in Dutch indicating that John is absent. Therefore, the expression is considered to be a grammatical expression: copula ‘be’ together with an infinitive form the grammatical ingredients, in a similar way as the combination of copula ‘be’ + aan het + infinitive form the grammatical expression of progressive in Dutch. The absentive contains four types of information: (i) the referent of the subject (=Subj) is absent, (ii) the Subj is involved in an activity, (iii) it is predictable how long the Subj will be absent, and (iv) the Subj will return after a period of time (De Groot 2000:695). One context in which the absentive is most naturally used is the

---

following question – answer pair (on the phone/at the door): A: “Could I speak with Peter, please”; B: No, sorry, he is off playing football”.

When we return to Old and Early Middle English, we see two constructions which form the basis for the later progressive, namely the construction with the present participle *he wæs huntende* and the construction with the verbal noun *he wæs on huntunge*. According to Bertinetto et al. (2000:530), there is some evidence suggesting that the first examples of progressive constructions could be used in contexts which exclude a purely focalized reading. Consider Old English (Bybee et al. 1994:135):


Some *syndan creopende* on eorða mid eallum lichoman, swa wurmas
doð. Some *gað* on twam fotum, some on feower fotum, some *fleoð* do.
some *go* on two feet some *fett* some *fly* with *fyðerum*.

'Some (animals) creep on the earth with their whole body, just as worms do.
Some walk on two feet, some on four, some fly with their wings'.

In contrast to the constructions with the present participle, I will show in the following section that there are good reasons to believe that the construction with *be* + *on* + verbal noun in Old and Early Middle English could have been a construction to express absence from the deictic centre.

3. Absentive properties of constructions with a verbal noun

In Ælfric’s Colloquy, we find the following sentence: *gyrstandæg ic wæs on huntunge* (p. 24, 1.68-69). The form is so rare in Old English that Åkerlund (1914:323), who never came across it, took it to be a ‘ghost-phrase’. In Middle English it is found more often, though by no means as frequently as *to be* + present participle (Scheffer 1974:244). In the following example we find a strong clue that the construction with the verbal noun is used to express the dislocation of the king from the deictic centre, i.e. the place where the dialogue takes place.


And whan *Pat hende knight Beuoun* 
Come wiP *outen Pe toun*,
Par wiP a palmer he mette,
And swiP faire he him grette:

'Palmer', a sede, 'what is Pe king ?'
'Sire !', a sede, 'an honting'

WiP kinges fifene.'
‘And *whar*, a seide, ‘is Pe quene?’ 2056
‘Sire’, a seide, ‘*in her bour*.’ 2057

Note that this fragment contains two questions introduced by *whar* ‘where’. The interrogatives inquire after the location of the king in (2053), and the queen in (2056). Instead of specifying a place by a noun and preposition, for instance ‘on Burch’ (in Peterborough), the first answer uses a verbal noun and preposition: *an honting* (2054). From this answer, the Addressee will derive the information that the king is out (not present) and somewhere out in the fields or forest, the place for hunting animals. The answer to the second question, however, specifies the place by a preposition and a noun phrase: *in her bour* (2057). In other words, the construction with a verbal noun can be used to specify a location in Middle English.

Example (9) contains a firm indication that the use of the verbal noun with the preposition *on/an* has a meaning of dislocation or absence. In the following subsections more evidence will be brought forward which supports the view that the construction with a verbal noun was indeed used to express somebody’s absence from the deictic centre.

### 3.1 *Esse in venatione* glossed by *be on huntunge*

The oldest example of a construction with the verbal noun is found in Ælfric’s *Colloquy* (p. 24, 1.67-69; 80). Ælfric wrote the *Colloquy* as an educational aid for teaching boys Latin. The text consists of a dialogue between a schoolmaster and his pupils. In addition to the Latin text, there is a gloss in Old English. A characteristic of Ælfric’s writing is that he was able to impart liveliness and spontaneity to the dialogue of his questionnaire. In the introduction to the text, the editor even states that ‘to-day the work is chiefly of interest for the picture it presents of the life and activities of the middle and lower classes of Anglo-Saxon society’ (Introduction, p. 1). The following fragment from the *Colloquy* follows on a conversation with the topic ‘things one is able to do: hunting’.

(10) *Old English* (Ælfric’s *Colloquy*, p. 24, 1.62-70)

Jea, butan nettum huntian ic mæg.

*Etiam sine retibus uenare possum*

Hu

*Quomodo*

Mid swiftum hundum ic betaec wildeor

*Cum uelocibus canibus insequor feras*

Hwilce wildeor swyÞost gefehst Þu ?

*Quales feras maxime capis*

Ic gefeo heortas 7 baras 7 rann 7 reegan 7 hwilon haran

*Capio ceruos et auros et dammas et capreos et aliquando lepores*
The context in which *on huntunge* in (68–69) is used does not require a progressive kind of expression. If, on the other hand, it could be established that *on huntunge* is a progressive, it could be a Durative progressive construction as mentioned in Section 2 above. However, in this example it is difficult to evaluate the action relative to a larger interval of time. The construction may also denote the absentive in the sense of ‘have you been out hunting?’ It is not possible to derive the exact interpretation from the context given. The Latin text, however, gives an interesting clue.

The Latin phrase which corresponds to *be on huntunge* is *esse in venatione*. Note that the form uses the preposition *in* and that the noun *venatio* is marked by the ablative case -ne. It is important to know that the form *esse in venatione* with the meaning of ‘being on a hunt’ does not occur in Classical Latin. The Latin used in the Colloquy is very similar to Classical Latin but it is a variant of Medieval Latin, or more specifically the Medieval Latin used in England. The author of the Colloquy was most probably not familiar with Classical Latin, but learned, as was quite common in his time, the Medieval variant. The meaning of *venatio* in Medieval Latin in England was different from Classical Latin. It did not mean ‘hunt’ or ‘hunting’, but ‘place for or right of hunting’ (Latham 1965:506). This explains why the locative expression with *in* and the ablative could be used, since the word referred to a place and not to an action as in

5. There are several alternative forms found in Old English texts, e.g. *huntunge*, *huntnolde*, *huntnoÞe*, and *huntnaÞe*.

6. No cases are attested in the Bibliotheca Teubneriane Latina (2002). I thank Rodie Risselada for her help.

7. Arie Wesseling, personal communication.
Classical Latin, where one should expect the use of the preposition *ad* + accusative as a direction towards. Given this information, I conclude that the form *esse in venatione* in the *Colloquy* was glossed by *be on huntunge* in the meaning of ‘be on a hunting place’, hence it was an expression of location. The use of the construction *be on huntunge* then easily fits in with an absentive interpretation.

3.2 The verbal noun was an abstract noun

Brunner (1962) shows that verbal nouns in *-ing/-ung* are formed with verbs of class 1 (*-ing*) and class 2 (*-ung*), of the weak verb classes. He argues that the verbal nouns are originally abstract nouns, as for instance in the following examples.

(11) *hit is . . . mæra on huntunge heorta and rāna*  
     ‘it is famous by its hunting of deer and reindeer’

(12) *sēo fēding Þēra scēapa*  
     ‘the feeding of the sheep’

On the basis of the following observations, Brunner (1962:253–258) illustrates how the verbal nouns acquire verbal properties in a later stage.

(13) a. Verbal nouns are used without an article.  
    b. Pronouns occur in the Oblique form in stead of the Genitive (possessive).  
    c. Adverbial and prepositional modifiers are in postposition. In the case of verbal nouns, these modifiers precede the noun.  
    d. Modifiers take the adverbial ending *-ly*.  
    e. The formation of past tense and passive constructions with the *-ing* form of *have or be* + the past participle of the verb.

The development of verbal nouns from abstract nouns to more verbal elements does not in itself say anything about an absentive reading of constructions with a verbal noun. However, the development sketched does not contradict the view that the construction with an abstract (verbal) noun was used to express somebody’s absence. Or even better, it may support the view. Evidence therefore is that nominal elements are prototypically used for spatial orientations, whereas verbal elements are not. When in the course of time the construction developed more verbal properties, the construction became less suitable for expressing absence, and more suitable for expressing ongoing activity. The meaning of the construction grew towards the one with *be* + present participle, i.e. a progressive meaning.

---

8. Scheffer (1974:245) says the following about these constructions: "Verbal nouns with *of* follow directly from the Old English form *ic was on huntunge*, especially from *faran on huntunge hranes* (*Ælfric’s Colloquy*, p. 30, l. 113/4). Here *huntunge* is not a verb and is therefore not followed by a direct object, but by an attributive adjunct in the genitive, which was later, owing to the analytical tendency of the language, replaced by an *of*-adjunct."
The spatial and temporal use of *on*

In his study on the prepositions *at, on, and in* in Old and Middle English, Lundskær-Nielsen (1993) shows that *on* in Old English had a predominantly spatial use. In later stages, the preposition *on* was used more and more in a temporal sense.

In the entries of the years 892–900 in the *Anglo-Saxon Chronicle* the preposition *on* occurs 86 times, 55 are in locative constructions, some in idiomatic and possessive constructions and 9 show figurative or abstract use. *On* is used 13 times in a temporal way. The examples with abstract use are of particular interest here. Lundskær-Nielsen (1993:93) says the following:

Some of the nine examples here have the general meaning of ‘being on the way’ to some (usually) unspecified place or simply ‘being out’ somewhere; the location is not as clearly defined as in the other categories defined above, so we are dealing with an extended spatial notion. Thus we find *on anne sib* ‘in one journey/voyage’ (893), *on fare* ‘on the way/march’ and *on hergaþ* ‘harrying, on a raid’ (twice) (894).

This fragment contains three very interesting pieces of information. The first one is that of the general meaning of ‘being on the way’, which means absence from the deictic centre. The second one is the ‘extended spatial notion’. Lundskær-Nielsen recognizes that the examples contain both spatial information (a specific place), and deictic information (the place is distinct from the deictic centre). The third one is that the last example, *on hergaþ*, counts as a perfect example of a construction with an absentive meaning: “be off harrying”, and “be of on a raid”.

In the entries for the years 1122–1154 in the *Peterborough Chronicle*, the highest number of occurrences of *on* (60 out of 157) fall under the heading of temporal usage. *On* in verbal complements is in the majority of cases used in a purely spatial sense; that is, the main verb of the clause expresses ‘location’ or ‘non-movement’, e.g. *þes feorðe gæges Þæræfter wæs se king Heanri on Roueceastre* ‘the fourth day after that King Henry was in Rochester’. However, examples such as *pa tune on Burch* ‘the town of Peterborough’ suggest that *on* could suffer a devaluation of meaning, which Lundskær-Nielsen considers early examples of grammaticalization of locative prepositions. Moreover, the preposition may also be used in directional complements, as in (14):9

(14)  Ond Þone eorl he sende to Briige on heftunge
     ‘and he sent the earl to Brigg into confinement.’

On the basis of the data presented by Lundskær-Nielsen, it is clear that in the years between 892 and 1154 the preposition *on* developed from a preposition with a predominant spatial use to a preposition with several usages, with the highest number of occurrences falling under a temporal use.

3.4 The combination with *ridan* and *owt*

Scheffer (1974:244) mentions that the analogical form with a verbal noun also occurs with other verbs than *to be*, e.g. with *to ride*.

(15) Middle English (King Horn, MS C, ed. E.E.T.S., p 29, 1. 684; MS L, p. 30, 1. 700; MS H)
   a. Pe king him rod an huntinge
   b. Pat ich rod on fischinge
   c. Pat ich rod ofysshynge

The examples in (15) have unmistakably an absensive meaning, conveying that “The king rode *out* hunting” and “I rode *out* fishing”.

The most convincing example illustrating the absensive nature of the construction with *be* + *on* + verbal noun is the following, mentioned by Scheffer (1974:245), with an overt expression of absence in the word *owt* ‘out’.

(16) Middle English (Stoner Letters and Papers; ed. Camden Society; 2. 123)

  John Cheynye is owt a hawking

  ‘John Cheynye is out on hawking’

Languages which have the Absentive sometimes show variation between expressions with and without “out”. Dutch is one of the languages where the variation between the two expressions, however, is not free.11

3.5 Marking Figure or Ground

In explaining the rise of certain periphrastic constructions, Kuteva (2001:51) introduces the notions of Figure and Ground. In the English sentence ‘The picture is on the wall’, for instance, ‘the picture’ is the Figure and ‘the wall’ the Ground. The linguistic expression of spatial configurations across languages involves a varying degree of precision regarding the information about the Figure and the Ground. The relation between objects in space specifies the geometrical nature of Ground, not the Figure. Ewe, for instance, is one of the languages which provide precise information about the nature and the dimension of the Ground: “The picture is on the wall/on the higher part of the wall/on the top of the wall”, etc. In languages where the Ground is presupposed, the Figure may be specified. This is for instance the case in languages such as Dutch, where postural verbs may be employed in progressive constructions, for instance *Jan*

10. The example already shows the reduction of *on* (an) to *o* (a). With ‘hunting’ the first reduction to *a* is found in the Chronicle of William Gregory 1470 (1. 219): *rode a huntinge* (Mossé 1938:Vol. II, p. 110).

The king is on huntunge

*zit te wachten* (John sits to wait 'John is waiting') and *Jan staat te wachten* (John stands to wait 'John is waiting'). Dutch, however, has both options to mark the Figure (17a) or the Ground (17b).

(17) a. Jan loopt te jagen. 'Figure'
    John walks to hunt
    'John is hunting.'

b. Jan is op jacht. 'Ground'
    John is on hunt
    'John is out hunting.'

Interestingly, where the Figure is marked, the clause has a progressive meaning, whereas if the Ground is marked, there is an absentive meaning. I do not claim that it will be always the case that an absentive meaning can be derived from constructions where the Ground is marked. They merely license an absentive interpretation.

English does not mark the Figure, neither in modern nor in earlier stages of the language. As for the Ground, let us return to the Old and Middle English constructions. No Ground can be established for the construction with *be* + present participle. There is just the subject and a predicate which consists of an auxiliary verb with a present participle. The second construction with *be* + on + verbal noun, however, contains a subject, a copula and a non-verbal, i.e. locative predicate. The construction has a Ground specified by the verbal noun, marked by the preposition *on*. The construction is spatial indeed, and may therefore have the absentive meaning. The difference between the two constructions in that Ground is not relevant to the former but relevant to the latter, supports the view that the two constructions contrast in the sense that the former has a progressive interpretation and the latter an absentive one. Progressives do not necessarily specify a Ground, whereas Absentive constructions most likely do.

3.6 Class of verbs

When we look at the verbs from which the verbal nouns are derived, we see that they constitute a homogeneous class. They all denote activities performed in particular places. The verbal nouns we find are for instance 'hunting,' 'hawking,' 'fishing,' 'harrying,' and 'fighting.' This again supports the view that the construction with verbal nouns was meant to express someone's absence from a deictic centre.

---

12. In fact there is a progressive construction in Dutch where the Ground is marked and not the Figure, as in *Jan is aan het boxen* [John is at the box,INF] 'John is boxing.' However, in present-day Dutch the preposition *aan* has lost most of its spatial usage.
3.7 Preliminary conclusions

In this section we have seen that there are at least seven different kinds of arguments that favour the idea that the construction with a verbal noun and a preposition in Old and Middle English was an absentive construction. The arguments taken individually do not make a strong claim; however, all arguments together present a very strong case. Briefly, these arguments are as follows:

(18) a. The construction is found in a dialogue where it is used as an answer to a where question.
    b. The construction is used as the translation of Latin esse in venatione.
    c. The construction is nominal and not verbal.
    d. The preposition has a spatial use.
    e. The construction combines with ridan ‘ride’ and owt ‘out’.
    f. The construction can be analysed as a construction in which a Ground is marked.
    g. The class of verbs denote highly specific activities performed in particular places.

As mentioned in Section 2 above, the absentive as used in a number of present-day languages of Europe contains four types of information: (i) the referent of the subject (=Subj) is absent, (ii) the Subj is involved in an activity, (iii) it is predictable how long the Subj will be absent, and (iv) the Subj will return after a period of time. The examples from Old and Middle English do not all show that they meet these requirements, so that they could count as the Absentive in the modern sense. However, when we consider the dialogue from the Romance of Sir Beues of Hamtoun, the example does fully fit the requirements.

(19) ‘Palmer’, a sede, ‘whar is Þe king ?’
    ‘Sire !’ a seide, ‘an honting WiÞ kinges fiftene.’

The pilgrim could have answered ‘not here’ or ‘in the countryside’, but he did not. By answering ‘an honting’ he offers the following types of information:

(20) a. The king is not present.
    b. Based on pragmatic knowledge, the Addressee will understand that the king will be away for some time (assuming that the king goes on a hunt for a couple of days or even longer).
    c. The king will return after the hunt.

When the Absentive is used in languages such as Dutch, all these types of information are contained in the expression.
4. Discussion

In the literature on the rise of the progressive in English, the construction with the present participle is usually taken to be the source construction. One other construction, the one with the verbal noun, began to develop verbal properties. It is suggested that the two constructions were semantically close and that in connection with phonological changes the following two separate Old English constructions he wæs huntende and he wæs on huntunge became very similar in Middle English *he was huntyng(e) and he was on/an/in/a huntyng(e), and ultimately coalesced, thus sharply increasing the frequency of the progressive form proper (cf. Fischer 1992). In this paper I make a case for analysing the construction with the verbal noun as an absentive construction. In Section 3, I have given seven reasons why I consider the construction with the verbal noun to be a construction denoting absence from the deictic centre.

Given the claim that the construction with a verbal noun is indeed an Absentive construction, it follows that the construction with the present participle (progressive) and the construction with the verbal noun (absentive) were not semantically close, as suggested in the linguistic literature, i.e. the construction with the verbal noun was not a (kind of) progressive too. Still, constructions denoting a progressive meaning on the one hand and constructions denoting an absentive meaning on the other hand may grow towards each other. The main reason for this is that the progressive and the absentive partially overlap. The progressive may in certain cases denote absence, e.g. in the case of “John is playing tennis”. If the deictic centre is not the tennis court, then in addition to the progressive there will also be an absentive reading. Although the absentive is basically a spatial deictic construction, the construction is, however, often felt to be a kind of progressive. Moreover, in Finnish there is a construction which can have both the progressive and the absentive interpretation. The interpretation of progressive or absentive in Finnish is determined by word order. Within a cognitive construction approach as used in Goldberg (2003) or Verhagen (this volume), I assume that the progressive construction and the absentive construction are closely related in a taxonomy of constructions. In other words, the semantic relation between progressive and absentive may be one of the reasons why the two distinct constructions in Middle English fused into one.

As for the second reason, note that many progressive constructions in the languages of the world find their origins in spatial constructions (cf. Bybee et al. 1994 and Krause 2002). A clear example of the spatial origin of the progressive is found in German: Sie ist beim Skatspielen [she is at skatplaying] ‘She is playing skat’ (Ebert

1996:46f.; Krause 2002:44f.). Given this line of development, the absentive, being a spatial construction, may under certain conditions easily undergo the same process. Evidence for such a development in English is the development of the preposition on from spatial in Old English to temporal in Middle English (cf. Section 3.3 above). The development could be schematically represented as follows:

\[(21) \quad \text{absventive} \quad \rightarrow \quad \text{progressive} \quad \rightarrow \quad \text{progressive construction} \quad \rightarrow \quad \text{the king is hunting} \]

On the basis of the data available, it is not possible to establish the exact path in which the two constructions melt together. There are two possibilities: (i) the construction with the present participle developed into the progressive; the absentive was reinterpreted as a progressive (cf. (21)); the two progressive constructions coalesced, or (ii) both constructions with the present participle and the verbal noun developed into the progressive simultaneously.

Nehls (1974) and Dal (1952) make explicit statements about the difference between the usage of the two constructions. According to Nehls the construction with the verbal noun became dialectal/non-standard. It would be interesting to investigate the dialects of English in which the construction with the verbal noun remained and see what have happened with the constructions and their meaning.

According to Dal the construction with the verbal noun was used in ordinary or sub-standard speech, while the construction with the present participle was used in educated speech. Dal may be right that the two constructions belong to two different modes of language use. These modes, however, are not educated versus ordinary or sub-standard speech, but rather epic or narrative versus dialogue style. Narratives can be structured around one or more time points, to which absolute and relative reference can be made by means of different grammatical tense distinctions. Aspect, which concerns the internal time structure of events, is also often related to the temporal structure of a text, e.g. 'When John entered the room, Mary was reading a book,' where the event of reading the book (expressed by the progressive) is located around the time point that John entered the room. In contrast, the absentive relates to a deictic centre, which is often (but not necessarily) determined on the basis of extra linguistic information, e.g. the place where a conversation takes place. That is why we find genuine examples of the constructions with a verbal noun used as an absentive in dialogues only, and that is why we do not find many examples of the construction with the verbal noun in Old English, because the number and size of texts with dialogues is highly restricted. This then explains why the form is very rare in Old English texts.

---

14 Note that the Absentive in German takes the following form: Sie ist skatspielen. [she is skatplaying], 'She is off playing skat.'
The king is on huntunge

5. Conclusions

In this paper I have shown that Old English in transition to Middle English had a construction with the properties of the Absentive. The view that there is a distinction between the construction with the present participle (progressive) and the construction with the verbal noun (absentive) sheds new light on the relation between the two constructions in Old and Middle English. It makes explicit what the semantic difference between the two constructions is. A partial overlap in the semantics of the progressive and the absentive on the one hand, and the fact that many progressives find their origins in spatial constructions, on the other, motivate the possibility that the two constructions may have coalesced. The distinction between progressive and absentive also explains why the construction with the verbal noun (absentive) is rare in Old English, because the form predominantly occurs in dialogues, a form which is rare in old texts.

From a theoretical point of view, the data and the analysis offer interesting new material for the study of aspect and Aktionsart, the relation between progressive and absentive, grammaticalization and the history of the English language.

References


**Sources**


PART II

The architecture of functional models
Mental context and the expression of terms within the English clause

An approach based on Functional Discourse Grammar*

John H. Connolly
Loughborough University

In order to achieve an integrated pragmatic model of the use of natural language, an account is needed of the interplay between grammar and context in the formulation and interpretation of utterances. The purpose of this chapter is to outline how such an account may be developed for Modern English. It is proposed that information which is part of the interpretation of an utterance, but which is not represented grammatically, needs to be represented at a properly articulated contextual level. If this is done, then it is possible not only to retain a parsimonious type of grammatical representation, but also to accommodate explicitly the contextually supplied contributions to the interpretations of utterances.

1. Introduction

In a number of his writings on Functional Grammar (FG), Lachlan Mackenzie maintains the position that grammatical representations of expressions should be restricted as far as possible to overtly expressed elements; see for instance Mackenzie (1995:336–342, 1998:267–269). In the spirit of such an approach, we can appreciate that whereas the clause in (1a), for example, would be represented as containing an actor and a goal (both being overt, i.e. present in the linguistic expression), on the other hand the

* A paper entitled “The treatment of English spatial prepositions in FG” was presented at the Eighth International Conference on Functional Grammar at the Free University, Amsterdam in 1998. I am grateful to fellow participants in that conference for useful comments on the paper. I also thank Miss Elizabeth M. Connolly, M.Phys., for her help concerning the application of coordinate geometry to linguistic description.
nominalization in (1b) would be accorded a representation that accommodated “the president” but not “the rebel leader”.

(1) a. The president was assassinated by the rebel leader.
    b. The assassination of the president.

(2) a. Who bought the flowers?
    b. Alice.

Again, the representation of the clause in (2a) would contain both an actor and a goal, but the elliptical reply in (2b) would be assigned a representation that included the actor only, and excluded the covert elements “bought” and “the flowers”; cf. also Dik (1997:328–330). These elements would be supplied not through the grammatical structure but, rather, from the context.

Dik (1997:1–4) took the view that a grammar should be seen as a component of an integrated model of a natural language user (NLU). This integrated view demands that an account should be given of the interplay between grammar and context in the formulation and interpretation of utterances by NLU’s as part of the discourse process. In order to provide such an account in an intelligible manner, information which is part of the interpretation of an utterance, but which is not represented grammatically, needs to be included within an explicit representation of the context. If this is done, then (as will be shown) not only can we keep the parsimonious type of grammatical representation favoured by Mackenzie, but we can also find a place to accommodate explicitly the contextually-supplied contributions to the interpretations of utterances.

In order to provide a proper account of context within FG, it is proposed in Connolly (2004) that the framework should include a contextual level, and that this level should therefore form part of the latest version of FG, namely Functional Discourse Grammar (FDG); see Hengeveld (2004). This proposal offers a ready-made home in which to accommodate the desired, explicit account of context, as will be demonstrated below.

The purpose of the present chapter is twofold. Firstly, an attempt will be made to show how the contextual level can be exploited in order to facilitate an integrated account of the formulation and interpretation of terms within the Modern English clause. Secondly, it will be suggested that the contextual level may provide an avenue towards the solution of an as yet unresolved issue in the formulation of, in particular, temporal and spatial satellite terms. The scope of the paper is entirely restricted to the FDG framework, as constraints of space preclude any useful discussion of alternative approaches.
Mental context and the expression of terms within the English clause

2. Terms, context and interaction

2.1 The classification of context

Let us begin with some general remarks about the context of language. The first point is that "context" is a relative term, in that it relates to (and, indeed, presupposes the existence of) some piece of language which represents our current focus of attention and which we shall therefore call the "focal segment". The context of that focal segment lies in whatever surrounds it.

Hence, context potentially constitutes a vast domain, and consequently, in order to handle it in a manageable way, we are required to take two major steps. Firstly, we need to impose an important constraint on the term "context". We stipulate that this term does not apply to everything in the universe that surrounds the focal segment, but only to what is relevant to its production and interpretation; cf. Nofsinger (1997:356). However, what is considered relevant is inevitably, at least in part, a matter of subjective judgment. This means that context is not an objective property of the universe but an interpretation either by a discourse participant or by a discourse analyst; cf. Shailor (1997:97).

Secondly, we need to classify context into categories. Various attempts have been made to do this; see for instance Firth (1957:203), Hymes (1972:58–65), Harris (1988:78–81), Devlin (1991:33, 217–221), Goodwin and Duranti (1992:6–9), Cook (1992:1–2) and Lyons (1995:271). In the light of such treatments, we may draw the following broad distinctions as an approach to the classification of context:

(3) a. Discoursal context versus situational context.
    b. Physical context versus socio-cultural context.
    c. Narrower context versus broader context.
    d. Mental context versus extra-mental context.

In order to assist with the explanation of these distinctions, let us take an example. Suppose a lecturer in a university department of Electronic Engineering is delivering a lecture, the first three clauses of which are as follows:

(4) a. In the previous lecture I mentioned that Hidden Markov Models are widely used in speech recognition technology.
    b. Today I want to explain how they can be used in written character recognition systems.
    c. So can you take a look at the diagram on the screen behind me? (While saying this, the lecturer points at the screen.)

Suppose, furthermore, that this discourse is being produced in circumstances in which all of the following facts are true:

(5) a. The participants in the discourse are:
    (i) The lecturer, who is the speaker.
(ii) A class of 30 students, who are the addressees.

b. The location is a room containing a screen visible to the addressees.

c. The time is 10.00 on Thursday, 29 April, 2004.

d. The social relationship of speaker to audience is that of teacher and learner.

e. The discourse is taking place within a university degree course.

f. The audience expects to be addressed in English.

The first distinction mentioned in (3) is that between “discoursal context” versus “situational context”. The discoursal context lies in the communicative activity surrounding (and relevant to) the focal segment. This communicative activity may include spoken or written language and also any non-verbal communication accompaniments, for instance meaningful gestures such as pointing. (This motivates the use of the term “discoursal context” rather than the narrower term “linguistic context”, given that discourse is in general multimodal; cf. Kress & van Leeuwen 1996:39.) For instance, the interpretation of the word they in (4b) as referring to Hidden Markov Models relies on the discoursal context furnished by (4a). The situational context, on the other hand, lies in the universe outside of the discourse, for instance the facts in (5).

The second contrast is that between the “physical context” and the “socio-cultural context”. The former encompasses phenomena that are part of the material universe, for example the physical location (such as a particular point on the earth’s surface) where a face-to-face discourse is conducted. Of course, if the speaker and addressee are in different places, then the physical setting of the discourse is geographically distributed over more than one location. The socio-cultural context, on the other hand, is associated with non-material factors such as interpersonal relations, social structures (including institutions, such as educational establishments) and cultural norms pertaining to beliefs, values, behaviour, and communication systems, including languages. Thus, (5a–c) pertain to the physical context, while (5d–f) belong to the socio-cultural context. (5a) enables the appropriate interpretation of the pronouns I/me and you; (5b) facilitates that of the phrase the screen and (5c) that of the word today; and (5d–f) account for the language and style employed.

The third distinction is between “narrower” and “broader” context, and it is applicable both to situational and to discoursal context. The narrower situational context includes the immediate location and time of the discourse; it also encompasses the participants involved in the communicative activity and their interrelationships, as well as any other relevant activity that is going on in the same setting, for instance a projector displaying an image on a screen. The broader situational context relates to wider circumstances, for example the identity of the official language of the country concerned.

The narrower discoursal context is found within the same communicative engagement as the focal segment, and has sometimes been characterized by means of the term “co-text”. Thus, as we have noted, (4a) provides a co-text for the interpretation of (4b). However, sometimes an “inter-textual” context arises. This happens when one dis-
course provides the context for another, as for instance in (4a–b), where the contents of the previous lecture assist the interpretation of the current one.

The fourth contrast is between “mental” and “extra-mental” context. Clearly, any aspect of context that is to influence the production or interpretation of discourse must be represented in the minds of the human NLUs concerned. These cognitive representations comprise the mental context, and normally some of them will be shared by speaker and addressees, even before the discourse begins. The entities and processes in the outside universe that are the objects of these representations belong to the category of extra-mental context.

When a discourse is in progress, each of the participants will have a memory of what has been said so far, and this memory constitutes part of the mental context. It may or may not include a faithful recollection of the actual words uttered, but even if it does, it will not be simply a mental representation of the preceding co-text. Rather, it will have been formed through an interpretation of the discourse, on the basis not only of the actual words uttered but also of other contextual factors, which together provide what Janssen (this volume) calls the ‘frame of reference’ for the interpretative process.

As the focus of attention moves, so the context will change. For instance, when (4b) is uttered, its immediate discoursal context is supplied by (4a). However, when (4c) is uttered, (4b) becomes contextual. This illustrates another important property of context, namely that it is not static but dynamic.

2.2 The contextual level

As stated earlier, it is proposed in Connolly (2004) that FDG should contain a contextual level. This level is not actually part of language-structure, but is, nevertheless, intended to facilitate the description of language by explicitly representing facts relevant to the production and interpretation of discourse. (It may be noted that Systemic Functional Grammar also recognizes a level of context; see Hasan (1999:224).)

Let us illustrate the use of the contextual level by means of a simple example. Suppose that Ann and Bob are chatting together in a bar, and Ann notices that Chris has just entered. She assumes that because Bob is facing away from Chris, he will not have seen him come in, but she wants Bob to know that he is there, lest Bob should make some derogatory remark about him. So she says:

(6) Chris is here.

Bob nods in reply, showing that he has understood.

A description of the above scene in terms of the contextual and interpersonal levels might run as shown in Figure 1. (The contextual information has been kept more-or-less to a minimum here, and more could be added if so desired.) The description states that the two discourse-participants (P₁ and P₂) are Ann and Chris, while Bob is not involved in the discourse, but is present in the role of a bystander (B₁), potentially able to overhear the conversation. Ann now decides that her next discourse-move (M₅₁)
Contextual Description

P₁ = Ann
P₂ = Chris
B₁ = Bob

Preconditions (M₅₁) =
KNOW (P₁ PRESENT (B₁))
BELIEVE (P₁
NOT (KNOW (P₂ PRESENT (B₁))))
WANT (P₁
KNOW (P₂ PRESENT (B₁)))

Interpersonal Description

ENTER MOVE M₅₁
Pₛₚ = P₁
Pₛₐₜₐ = P₂
M₅₁: [CONVEY (Pₛₚ) (m₅₁) (Pₛₐₜₐ)] (M₅₁)
ENTER ACT A₇₅
A₇₅: [ASSERT (Pₛₚ) (m₅₁) (Pₛₐₜₐ)] (A₇₅)
(details omitted for brevity)
LEAVE ACT A₇₅
LEAVE MOVE M₅₁

Postconditions (M₅₁) =
KNOW (P₂ PRESENT (B₁))
BELIEVE (P₁
KNOW (P₂ PRESENT (B₁)))
NOT (WANT (P₁
KNOW (P₂ PRESENT (B₁))))
Outcome(M₅₁) = Postconditions(M₅₁)

Figure 1. Description of a simple discourse in terms of the contextual and interpersonal levels

will be motivated by three considerations. These are known as the preconditions of the planned discourse-move, and they are as follows:

(7) a. Ann knows that Bob is present.
b. Ann believes that Chris does not know that Bob is present.
c. Ann wants Chris to know that Bob is present.

Accordingly, she makes the discourse-move, by saying:

(8) Bob's here.

In this move she is, of course, the speaker (Pₛₚ), while Chris is the addressee (Pₛₐₜₐ), and her intention is to convey to him the message (m₅₁) expressed in her utterance (8). The move comprises just one act (A₇₅), in which she asserts the content of that message. The fulfillment of her intention lies in the satisfaction of the three postconditions:

(9) a. Chris knows that Bob is present.
b. Ann believes (correctly) that Chris knows that Bob is present.

c. Ann does not (any longer) want Chris to know that Bob is present (since he is, by this time, already aware of the fact).

As Chris duly understands Ann’s utterance as intended, the outcome is that the intended postconditions do, indeed, hold true.

2.3 The formulation and interpretation of terms

Now that we have seen an example of how the contextual level may be described in tandem with a linguistic description of the discourse process, we may explore its use in more detail. As this chapter is concerned with mental context, we shall confine our treatment to this, and focus especially on the preconditions that motivate particular linguistic expressions. The treatment will be linguistically orientated, and no attempt will be made to elaborate upon the psycholinguistic details of mental operations. However, we shall feel free to make use of some concepts from set theory.

The question of what might constitute an appropriate kind of representation for the description of mental context also raises a problem, as insufficient is known about how people conceive mentally of situations to provide us with a definitive guide. In this chapter, we shall simply seek to employ representations, inspired by Vet (1998), that help to clarify the interaction between language and context, and we shall make no claim of psychological reality in matters of detail. Capitalized words will be used to denote the mental concepts conventionally expressed by the linguistic items concerned. However, this paper will not be concerned with the finer details of notation.

One possible use of the contextual level lies in handling situations in which the same facts may be linguistically expressed in a variety of ways. For instance, suppose that Ann knows the following facts:

(10) a. Derek has applied to Chris for promotion, and the time has come to make the decision. Ann is present with Chris, but Derek is not.

b. Chris rejects the application.

c. The time is 11.00 on Thursday, 29 April, 2004.

d. The place is room A2 at the headquarters of the British Mudguard Company in Staines, England.

It is Ann’s job to communicate the decision to Derek. As she composes the letter, the facts in (10) are part of her mental context. She writes a letter containing the sentence:

(11) Your application has been rejected.

She could, of course, alternatively have written:

(12) Chris has rejected your application.

However, she did not wish to attribute the decision to Chris, as she felt that to do so would have personalized the decision in such a way as to make future dealings be-
tween Derek and Chris more difficult. In other words, her mental context contained the following:

(13) a. She wanted Derek to know that his application had been rejected:
    b. WANT (ANN KNOW (DEREK REJECT (SOMEONE APPLICATION)))

(14) a. She did not want Derek to know that Chris had rejected the application:
    b. NOT (WANT (ANN KNOW (DEREK REJECT (CHRIS APPLICATION))))

Nor did she want to mention the precise time or place, as they were presumably of no interest to Derek. Nevertheless, she did have the option of including this information if she had so desired.

The various alternatives available to Ann arise because she has a mental representation of the facts in (10b–d) that may be characterized as follows:

(15) a. Action: REJECT.
    b. Performer of action: CHRIS.
    c. Affected by action: APPLICATION.
    d. Location: ROOM A2, BRITISH MUDGARD COMPANY, STAINES.
    e. Time: 11.00, 29/04/2004.

Let us regard each of (15a–e) as an “item of information”. (For present purposes, an item of information can be thought of as the content expressible by a single, simple element of the clause, where “simple” means “not containing any embedded clause”. This will provide a sufficient basis for what follows.) In formulating her sentence, Ann had the task of partitioning the set of five items of information into two subsets:

(16) a. A subset to be included in the message: (15a, c).
    b. A subset to be omitted from the message: (15b, d–e).

As demanded by the FDG approach, the excluded items never enter into the linguistic formulation, and therefore never have to be deleted if unexpressed.

In such a framework the production of an utterance would, thus, be modelled as follows:

(17) a. Begin with a mental contextual representation of a situation, in the form of a set of items of information.
    b. Partition this into two complementary subsets:
       (i) An “overt” subset, to be expressed as part of the message.
       (ii) A “covert” subset, to remain unexpressed.
       (This has the effect of thus filtering out the information in the covert subset. However, the filtering takes place before any involvement of the linguistic apparatus.)
    c. Attempt to formulate an expression to communicate the information in the overt subset only.
Because it starts up from a contextual footing, such a model is strongly pragmatic in character, as befits a functionally-oriented theory like FDG.

This model implies the existence of rules applying to the different stages in the overall process. Firstly there need to be rules governing the partitioning procedure. Very informally, these might be drawn up along the following lines:

\begin{enumerate}
\item If you believe that for social reasons it is better not to identify the performer of an action, then assign the performer to the covert subset.
\item If you believe that an item of information is of no interest to the addressee, then assign it to the covert subset.
\item If you believe that an item of information is already known to the addressee, then assign it to the covert subset.
\item Any item not assigned to the covert subset should be assigned to the overt subset.
\end{enumerate}

However, much more work would clearly need to be done in order to draw up a comprehensive set of rules that gave rise only to grammatically expressible, correct output. For a discussion of the problem of expressibility, see Bakker (1994:180–186).

Secondly, there need to be rules governing the formulation of the linguistic expressions. These are part of the grammar and are therefore already to some extent familiar. On the other hand, the partitioning rules lie outside the grammar and should be regarded, rather, as rules pertaining to verbal interaction, or 'pragmatic rules'; see Dik (1997:3–4) and Connolly (forthcoming).

Rule (18c) leads us back to the topic of ellipsis. For the purpose of illustrating the applicability of the contextual level here, we shall take a comparatively straightforward example, while acknowledging that not all cases are so simple. Nevertheless, our example is of a type that occurs quite frequently in practice.

Consider, then, a dialogue containing the following exchange:

\begin{enumerate}
\item Who rejected the application?
\item Chris.
\end{enumerate}

The first speaker’s mental context that lies behind the expressed question (19a) may be represented as follows, with the performer of the action being unknown and therefore left blank:

\begin{enumerate}
\item Action: REJECT
\item Performer of action:
\item Affected by action: APPLICATION
\item Location: not of interest
\item Time: not of interest
\end{enumerate}

Assuming that the second speaker knows the answer, then his mental context contains at least (15a–c). However, rule (18c) would cause the information in (15a) and (15c) to be assigned to the covert subset, leaving rule (18d) to assign only (15b) to the overt subset. This, of course, is exactly what is needed in order to produce the elliptical
response (19b), whose underlying linguistic structure may be represented along the following lines:

\[(21) \text{ (RP12): \[(d1x1: \text{Chris}(x1))\] (RP12)}_{\text{Agent}}\]

\(\text{RP} = \text{referential phrase}\)

When the second speaker receives the reply, his task is to interpret it in the light of the context. Given that his mental representation of the context includes (20), his task amounts to unifying the information conveyed by the utterance (19b/21) with that already present in (20), to produce the following mental context:

\[(22) \text{ a. Action: REJECT.} \]
\[(22) \text{ b. Performer of action: CHRIS.} \]
\[(22) \text{ c. Affected by action: APPLICATION.} \]
\[(22) \text{ d. Location: not of interest.} \]
\[(22) \text{ e. Time: not of interest.} \]

Thus, by including this contextual representation within our description, we are able to accommodate as full a specification of the meaning of the elliptical utterance (19b) as we desire, while allowing the underlying representation of the linguistic expression itself to be as parsimonious as we wish. In this way, then, we can have our cake and eat it too!

3. Temporal and spatial satellite terms

3.1 The expression of adpositional terms

Let us now consider the issue of time and location. If these dimensions figure overtly in a linguistic expression, then they may be expressed in a variety of ways. There is no space here to cover all the possibilities. Rather, we shall concentrate on one of the most frequently-used alternatives in English, namely expression by satellite terms in which an adposition is employed to indicate the temporal or locative meaning. Such adpositional terms are of particular interest in the present volume because they have received a novel treatment by Lachlan Mackenzie. We shall return to Mackenzie’s contribution shortly, but first of all we need to wind the clock back, for a moment, to the initial formulation of FG in Dik (1978).

3.2 A first alternative

In the original formulation of FG, only nominal, adjectival and verbal predicates were regarded as parts of the lexicon; see Dik (1978:15). All other classes of item were introduced into linguistic expressions by means of rules, and in particular, adpositions were inserted, where appropriate, through the application of expression rules; see
Dik (1978:20, 157). Adpositions were thus seen as comparatively superficial signals of underlying semantic functions.

In Connolly (1994, 1995) an attempt was made to formulate a set of expression rules for the purpose of inserting adpositions into English temporal expressions, assuming the framework proposed by Dik (1978). However, in order to produce a viable set of rules, it proved essential to expand the underlying representations of temporal terms beyond what had previously been considered necessary, and to include within some (though not all) of those underlying representations what are called “relational formulæ”, along the lines of the following example:

\[(23) \ t_1 < t_2\]

This indicates that a certain point in time \(t_1\) is earlier than another point in time \(t_2\). Thus, for instance, the temporal term in (24a) would have an underlying representation along the lines of (24b):

\[(24)\]

a. [The letter will arrive] before noon.

b. .. (.. (d1t_1; (d1t_2; noonN(x_2): t_1 < t_2)) ..)Time-position ..

where \(t_1\) refers to the time of the “arrival” event (\(e_1\)).

(Irrelevant parts of the notation are omitted here and in what follows. However, acknowledgement must be given to François (1996), whose discussion has led to improvements incorporated in the present chapter.) The rule in (25) for inserting the adposition \textit{before} would then rely on the presence of the appropriate relational formula in the underlying representation of the term concerned:

\[(25) \ Time-position [\text{term}] = \text{before} \text{term} \]

\[(\text{if } t_j \text{ is the time-position of } e_1 \text{ and } t_i \text{ is the time-position of } \text{term} \text{ and } t_j < t_i)\]

Using this approach (with some additions which need not detain us here) a reasonably comprehensive set of rules was formulated for temporal terms.

Furthermore, it is possible in principle to apply the same kind of approach to spatial terms. For this purpose, of course, we need to recognize not just the one dimension required for the treatment of the time axis, but three dimensions. These may be represented in terms of three mutually perpendicular axes labelled X, Y and Z, forming a system of Cartesian co-ordinates with origin O. As illustrated in Figure 2, any point P within this three dimensional space may be identified by the set of co-ordinates \((x,y,z)\), denoting the distance of P from O along the X, Y and Z axes respectively.

It is assumed that the X axis is horizontal, running from left to right, the Y axis is vertical, and the Z axis represents depth in the scene, with nearer points having a higher value than further points. For an earlier use of co-ordinates in FG, see Marshall (1993).

In addition, the following shorthand notations will be used in what follows:

\[(26)\]

a. \(\int A = \) every point within or on the boundary of object A.

\(\int B = \) every point within or on the boundary of object B.
Figure 2. A point defined in terms of Cartesian co-ordinates

\[ P(x, y, z) \]

b. \( y(\int A) \leq y(\int B) \) means: in terms of the (vertical) Y axis, every point within or on the boundary of object A is below or level with every point within or on the boundary of object B. (If at least one point on the boundary of A is level with at least one point on the boundary of B, then this simply implies that A and B are in contact.)

To take an example, the spatial term in (27a) would have an underlying representation along the lines of (27b):

(27) a. [The potato is] under the cooker.
   b. .. (.. (d1p1: (d1x2: cookerN(x2): y(\int x1) \leq y(\int x2))) ..)Locative ..
   where \( x1 \) designates the "potato".

The rule for inserting under into the linguistic expression would then be as follows:

(28) Locative [term] = under term
    if \( p_i \) is the location of \( x_j \) and term \( x_i \) is the point of reference and if, for the referents indexed by \( x_i \) and \( x_j \):
    \[ y(\int x_i) \leq y(\int x_j). \]

This rule applies to the term \( (x_i) \) representing the "point of reference", namely the location to which the speaker intends to relate the antecedent term \( (x_j) \) whose spatial position is being described by the locative term \( (p_i) \) within which \( (x_j) \) is embedded. Thus, in (27) the cooker represents the point of reference, and forms part of the locative term under the cooker, which describes the spatial position of the antecedent term the potato.

Again, various additions would be needed to handle all the spatial adpositions. However, the illustration just given serves to convey the basis of the approach.
3.3 A second alternative

In contrast to the approach just described, Mackenzie (1992) treats the spatial adpositions in a rather different manner. He proposes that only some of these should be inserted by means of expression rules, while others should be incorporated into the lexicon as adpositional predicates. This proposal represents a significant innovation to the framework of Dik (1978).

Mackenzie (1992:6) proposes five semantic functions for spatial terms, each of which may be directly expressed by a particular adposition, as shown in (29):

\[(29)\]
\begin{align*}
\text{a. Locative:} & \quad \textit{at} \\
\text{b. Source:} & \quad \textit{from} \\
\text{c. Path:} & \quad \textit{via} \\
\text{d. Allative:} & \quad \textit{to} \\
\text{e. Approach:} & \quad \textit{towards}
\end{align*}

He terms the five adpositions in (29) “grammatical prepositions”. All the other spatial adpositions he classifies as lexical. For instance, the spatial term in (30a) containing the adposition \textit{under} would have an underlying representation along the lines of (30b):

\[(30)\]
\begin{align*}
\text{a. } \text{[The saucer is] under the cup.} \\
\text{b. } \ldots (\ldots (\text{d1p: underP(d1x2: cup(x2))Reference }) \ldots )\text{Locative } \ldots
\end{align*}

This adposition would, of course, be introduced into the clause during the process of term formation.

Mackenzie's approach makes for less complexity in the underlying representations than is characteristic of the first approach described above. However, as François (1996:8) remarks, this complexity will not simply disappear, but will need, instead, to be handled in some other module within the overall framework.

3.4 A third alternative

Could the module in question perhaps be the contextual level of FGD? This is an idea well worth exploring, as it would offer a possible means of reconciling the two alternative approaches just outlined, in such a way that the underlying representation of terms would be spared the complexity attendant upon relational formulae, while at the same time a place would be found for incorporating into the linguistic description the analytical information which they help to make explicit. The idea is justified insofar as speakers can be presumed to form a mental conception of spatial relationships prior to describing them in language.

Again, let us work on the basis of an example. Suppose that the following facts are true in Ann's mental context:

\[(31)\]
\begin{align*}
\text{a. } & \text{Ann has seen that Bill has just dropped some vegetables, including a potato, on the floor of the kitchen and is trying to retrieve them.}
\end{align*}
b. Ann has seen that the potato has rolled under the cooker.
c. Ann knows that the potato is under the cooker.
d. Ann believes that Bill does not know that the potato is under the cooker.
e. Ann wants Bill to know that the potato is under the cooker.

Accordingly, Ann formulates an utterance to tell Bill. In so doing, she must decide which items of information to assign to the overt subset and which to the covert subset. For instance, the fact that the cooker is in the kitchen can be presumed not to be of interest to Bill, as he will know that already. Accordingly, she decides to mention only the potato and its location relative to the cooker, and says:

(32) The potato is under the cooker.

In order to describe the process of formulating (32) on the basis of a contextual-level description together with a set of rules, let us first suggest that the part of the mental context in (31e) be represented as follows:

(33) WANT (ANN KNOW (BILL (y(\int POTATO) \leq y(\int COOKER))))

The formulation of the underlying representation expressed by utterance (32) involves, inter alia, the following:

(34) a. Indexing POTATO as x_1 and COOKER as x_2.
b. Specifying x_1 as definite singular entity denoted by the appropriate lexical item: potatoN.
c. Specifying x_2 as definite singular entity denoted by the appropriate lexical item: cookerN.
d. Specifying p_1 as a definite singular location of x_2.
e. Inserting the adposition appropriate to the precise nature of the locative relationship between x_1 and x_2: underP.
f. Inserting the complement of the adposition: the x_2 term.

(Other parts of the process, such as the insertion of the x_1 term, are omitted for brevity.) In order to carry out (34e–f), we may adduce a rule along the following lines.

(35) composition[p_i] = p_i: underP (\ldots (x_i))_{Reference}
if p_i is the location of x_i and term x_i is the point of reference and the mental context indicates:
\gamma(\int x_i) \leq \gamma(\int x_i).

Here, \textit{composition}[p_i] means “the underlying structure and content of the p_i term”.

Examples of rules to introduce other lexical spatial adpositions would be as follows:

(36) a. composition[p_i] = p_i: behindP (\ldots (x_i))_{Reference}
if p_i is the location of x_i and term x_i is the point of reference and the mental context indicates:
Mental context and the expression of terms within the English clause

\[ z(\int x_j) \leq z(\int x_i). \]

b. \( \text{composition}[p_i] = p_i; \text{to\_the\_left\_of}(\ldots(x_i)) \_\text{Reference} \)

if \( p_i \) is the location of \( x_i \) and \( \text{term} \ x_i \) is the point of reference and the mental context indicates:

\[ x(\int x_j) \leq x(\int x_i). \]

To cover the insertion of all of the spatial adpositions would require various additional devices, and would also involve a discussion of non-literal uses of these forms; see further Herskovitz (1986) and Tyler and Evans (2003). However, the general idea can be gleaned from the examples provided.

What emerges is that the use of the contextual level makes possible a treatment of spatial adpositions that avoids the complexity arising from the inclusion of relational formulae in the underlying linguistic representations, while nevertheless accommodating those formulae, appropriately, within a description of the associated mental context and thus not losing the information contained in them. In this way, then, once more, we can have our cake and eat it too!

4. Conclusion

In this chapter some proposals have been presented which, it is hoped, may advance FDG in a direction consonant with the approach associated with Lachlan Mackenzie, through his distinguished research in this field. Our main conclusion is, in brief, that the inclusion of a contextual level within the FDG framework offers the interesting possibility of resolving some important and long-standing problems of linguistic description.

References


Adverbial conjunctions in Functional Discourse Grammar*

Kees Hengeveld and Gerry Wanders
University of Amsterdam

This contribution presents a new view on the treatment of adverbial conjunctions in Functional Discourse Grammar, combining an analysis of general linguistic interest with a theoretically motivated one. It starts off with a typological classification of linguistic elements expressing relations between clauses, which leads to a classification of conjunctions and conjunctional phrases in English. Attention is then paid to the different lexical-grammatical properties of different types of conjunctional elements. The findings show that, in English, linguistic elements used to combine clauses form a continuum ranging from grammatical elements parallel to prepositions to lexical elements parallel to verb chain contructions. The descriptive tools of Functional Discourse Grammar help to reveal a correlation between the types of conjunction and their domain of application.

1. Introduction

The aim of this paper is to investigate how adverbial conjunctions may be dealt with in Functional Discourse Grammar (Mackenzie & Gómez-González Eds. 2004; Hengeveld & Mackenzie forthc.). In doing so we concentrate both on the way the conjunctions themselves are represented in underlying structure, and on the nature of the linguistic units they conjoin. We furthermore intend to show that there is a correlation between the nature of the conjunction on the one hand, and the hierarchical level of the units connected by means of that conjunction on the other. We illustrate our theoretical points using mainly English examples, but do so only after sketching the general typological background within which the English examples may be situated.

The structure of this paper is as follows. In order to arrive at a proper delimitation of the notion of adverbial conjunction, we will in Section 2 first, from a typological

* We are grateful to Lachlan Mackenzie, who, unaware of the final destination of this paper, gave us a series of very helpful comments.
perspective, set them off from other linguistic elements expressing relations between a main clause and a dependent linguistic unit, such as special verb inflection and case markers. This section ends with a discussion of the types of conjunction available in English. On the basis of the typological classification we note that English draws a distinction between grammatical conjunctions (e.g. *because*), lexical conjunctions (e.g. *after*), and periphrastic conjunctions (e.g. *in the event*), which is reflected in the possibilities of modification of the various types. After briefly introducing the relevant parts of Functional Discourse Grammar in Section 3, Section 4 then discusses the underlying semantic representation of the various types of conjunction distinguished in Section 2. Section 5 then extends the analysis to the argumentative functions of conjunctions, making use of the distinction that is made in FDG between an interpersonal and a representational level of representation. We present our conclusions in Section 6.

2. Adverbial conjunctions

2.1 Adverbial clauses

Adverbial clauses are dependent clauses that are optional additions to a main clause, which means that they can be left out without affecting the grammaticality of the main clause with which they combine.¹ A defining feature of adverbial clauses is furthermore that they contain a specific marker that is indicative of the semantic-functional relationship between the main and the dependent clause (see Wanders forthc.). These two points may be illustrated by means of the following examples:

(1) Jenny closed the door *after* John left.
(2) Jenny regretted *that* John had left.
(3) *Coming out, stopping to check the mailbox, taking a look at the driveway and pausing to adjust his hat,* he turned and walked to his car.

The italicized units in (1)–(3) are all dependent clauses, but only the italicized unit in (1) is an adverbial clause. The italicized clause in (2), being a complement clause, cannot be left out without affecting the grammaticality of the construction as a whole, and therefore does not count as an adverbial clause. The italicized clauses in (3), taken from Givón (1995, see also Wanders forthc.), are part of a co-subordinating narrative construction. They can be left out without affecting the grammaticality of the main clause, but do not contain a specific marker that is indicative of a semantic-functional relationship with the main clause, and for that reason do not count as adverbial clauses.

¹. Note that according to this definition most conditional clauses and many correlative constructions (e.g. *The more I think about it, the more I disagree with you.*) do not count as adverbial clauses. These constructions too are probably better dealt with as cases of co-subordination (see van der Auwera 1997).
2.2 Adverbial conjoining

The marker that is indicative of the semantic-functional relationship between a main clause and an adverbial clause may take various forms across languages. The major difference that may be observed is the one between bound and free expression markers of the semantic-functional relationship. The following examples illustrate the bound strategy:

(4) **Evenki** (Nedjalkov 1997:53)

Evenki **bi-mi** nungan homoty-va davdy-ra-n

Strong COP-CAUS he bear-ACC.DEF win-REAL.3SG

‘He overcame the bear because he was strong.’

(5) **Basque** (Saltarelli 1988:47)

Ama-**k** errierta-Ø ema-n dio lorontzi-rik

mother-ARG.SG argument-ABS give-PF AUX.PRES flower.pot-PRT

on-en-a hauts-i bait-du.

good-MOST-SG.ABS break-PF CAUS-AUX.PRES

‘Mother has quarreled with him because he has broken her best flower pot.’

(6) **Imbabura Quechua** (Cole 1982:64)

Ńuka wawki shamus-shka-**manda-mi** jatun fishta-ta rura-rka-ni

my brother come-NR-ABL-FOC big party-ACC make-PAST-1

‘Because my brother came, I gave a big party.’

Evenki has a large set of specialized non-finite verb forms that may be used in adverbial clauses of various types. Thus, it has specific non-finite verb forms for Anteriority, Posteriority, Simultaneity, Condition, Purpose, etc. These non-finite verb forms are often referred to as adverbial participles or converbs (see Haspelmath & König 1995). In (4) the converbal ending -**mi** expresses causality. Basque uses affixes on finite verb forms to express certain semantic relations, among them causality, expressed by means of the verbal prefix **bait**- in (5). Imbabura Quechua in many cases nominalizes the adverbial clause and attaches a case marker to it. In (6) the ablative case marker expresses the causal relationship between main and nominalized clause.

There are a large number of varieties of the free strategy markers too. Consider the following examples:

**Mokilese** (Harrison 1976:260)

(7) Ngoah suh-**oang** John **anjoau-o** ma ngoah in-la sidow-a.

I meet-all John time-REM REL I go-DIR store-DEF

‘I met John when I went to the store.’

‘I met John the time at which I went to the store.’

(8) Ih dupukk-**oang** ngoahi **mwoh-n** oai japahl-do Mwoakilloa.

he pay-all I front-POSS my return-DIR Mokil
‘He paid me before I returned to Mokil.’
‘He paid me front of my returning to Mokil.’

(9) Kashmiri (Wali & Koul 1997:73)
Tik’azi sirinagœ as sakh garmi, amikin’ goœ bagulmarag.

Because Srinagar.ABL was very hot, therefore went I Gulmarg

‘Because it was hot in Srinagar I went to Gulmarg.’

(10) Spanish
Las calle-s están mojad-as porque está
def.pl street-pl cop.pres.3.pl wet-f.pl because cop.pres.3.sg
llov-iendo.

‘The streets are wet because it is raining.’

In the examples from Mokilese, the temporal adjuncts are actually noun phrases with a temporal noun as their head. These noun phrases furthermore contain a subordinate clause expressing the event with respect to which the main clause event is situated in time. In (7) the subordinate clause is a relative clause modifier of the temporal head noun anjoau ‘time’, and in (8) it is the second argument of the relational noun mwoh ‘front’. Strictly speaking, these are not cases of adverbial subordination, since the subordinate clause either modifies or is an argument of a noun. However, since this type of construction is often the diachronic source of true conjunctions, we discuss it here. Kashmiri uses a correlative construction: both clauses may be marked for the adverbial relation they are in, the relational element in the main clause (amikin’ ‘therefore’ in (9)) referring anaphorically to the event described in the subordinate clause. Finally, in the Spanish example in (10) only the dependent clause is marked, in this case through a grammaticalized adverbial conjunction.

2.3 Conjunctions and adpositions

It is noteworthy that in many languages the class of markers that are indicative of the semantic-functional relationship between clauses in adverbial constructions partly overlaps or shows similarities with the class of markers that are indicative of the semantic-functional relationship between noun phrases and their linguistic context. This holds for both bound and free markers. Consider the following examples:

Imbabura Quechua (Cole 1982:64, 116)


my brother come-nr-abl-foç big party-acc make-past-1

‘Because my brother came, I gave a big party.’

(12) Pay-ka chugri-manda wañu-rka-ø

he-top wound-abl die-past-3

‘He died because of his wound.’
2.4 Conjunctions and conjunctural phrases in English

If one accepts that the cosubordinate constructions in (3) are not adverbial clauses, it turns out that English uses free markers only to explicitly express adverbial relations. At least the following types may be distinguished:

(15) She called him before she left.
(16) She stayed home until the meeting began.
(17) Smallpox would be rapidly controlled in the event that it were introduced into Australia.
(18) I’ll bring him some water in case he gets thirsty.

The elements in italics in these examples can be classified along two parameters: simple versus complex, and lexical versus grammatical, as indicated in Table 1.

The subdivision between simple and complex conjunctions does not need further comments, at least for these examples. A distinction that is more crucial here is the one between lexical and grammatical conjunctions. A major feature that distinguishes lex-

<table>
<thead>
<tr>
<th></th>
<th>Lexical</th>
<th>Grammatical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>before</td>
<td>until</td>
</tr>
<tr>
<td>Complex</td>
<td>in the event that</td>
<td>in case</td>
</tr>
</tbody>
</table>

Table 1. Classification of conjunctions according to complexity and type
lexical conjunctions from grammatical ones is the fact that (parts of) lexical conjunctions can be modified through additional lexical means. Consider the following examples:

(19) She called him three hours before she left.
(20) *She stayed home three hours until the meeting began.
(21) In the unlikely event that smallpox were introduced into Australia, it would be rapidly controlled.
(22) *I’ll bring him some water in unlikely case he gets thirsty.

In (19) three hours gives a further specification of the time lapse preceding the occurrence of the event described in the dependent clause. In (21) unlikely qualifies the hypothesized event in terms of its reality status. Similar modifications of (part of) the conjunction are impossible in the case of until and in case.

The modifiers in (19) and (21) modify the conjunction (19) or part of the conjunctioanl phrase (21) directly, i.e. they have scope over a head. This is not the case with the modifiers in (23)–(26), which (may3) have phrasal scope:

(23) He arrived exactly three hours before she left.
(24) He continued walking around exactly until the meeting began.
(25) Only in the unlikely event that you don’t reply to this message will I phone you.
(26) Only in case it rains will I stay home.

The phrasal scope of the modifiers in italics in (23)–(26) can be brought out by replacing the entire subordinate clause by a simple anaphorical phrase, as in:

(27) exactly three hours before she left
> exactly then/at that moment
(28) exactly until the meeting began
> exactly then/at that moment
(29) only in the unlikely event that you don’t reply to this message
> only then/in those circumstances
(30) only in case it rains
> only then/in those circumstances

2. Although the distinction lexical-grammatical is a continuum rather than a dichotomy, the FDG model makes a sharp distinction between the two. Our position on this matter is that if modification of the meaning of (part of) the conjunction is impossible it is considered grammatical.

3. In (23) exactly may have narrow scope as well, if it is used to express the exactness of the time-span defined by three hours.
What may be deduced from these facts is that for instance *three hours before* in (27) is a single complex description of a moment in time based on a conjunctural phrase with a lexical head that can be modified directly. The complex description as a whole forms a (temporal) phrase that may be modified by phrasal modifiers.

A further consequence of the wider scope of phrasal modifiers is that they cannot be inside the scope of the head modifiers or occupy the slot of a lexical modifier, as illustrated in (31)–(32) for just two permutations, which are ungrammatical in the intended readings:

(31) *three hours exactly before she left
(32) *in the only event that you don’t reply to this message

Next to modification, there is a further phenomenon that points up the difference between a lexical and a grammatical conjunction, and which has to do with the combinability of lexical and grammatical conjunctions. Consider the following examples:

(33) She stayed until three hours after he left.
(34) She didn’t leave until the very moment he arrived.

In (33) and (34) *three hours after he left* and *the very moment he arrived* are complex descriptions of points in time, and *until* defines the time span leading up to those points in time. The opposite ordering of grammatical and lexical conjunctions is excluded. This combinability of grammatical and lexical conjunctions seems for semantic reasons to be restricted to temporal conjunctions, both simple (33) and complex (34), which are in this respect similar to grammatical and lexical locative prepositions, as discussed in Mackenzie (1992a, 1992b).

On the basis of the above considerations, we conclude that the grammar of English reflects a basic difference between lexical and grammatical conjunctions. In the next section we will provide a description of these two classes of conjunctions and their subclasses within the framework of Functional Discourse Grammar.

3. Functional Discourse Grammar

3.1 Introduction

Functional Discourse Grammar (FDG) as presented in (Hengeveld 2004a, 2004b, 2005; Mackenzie & Gómez-González 2004; Hengeveld & Mackenzie forthc.) is the grammatical component of a wider theory of verbal interaction (see Dik 1997), in which it interacts with non-linguistic components of the process of human communication. In the FDG model four interacting levels of organization are distinguished: the interpersonal level, the representational level, the expression level and the phonological level, in exactly that hierarchical order. It is characteristic of FDG that these levels are simultaneously present, i.e. where relevant linguistic units are fully analysed at each
of these levels. Internally, each of the levels of linguistic organization is structured hierarchically. Since the purpose of this contribution is not to explain the FDG model as a whole, we will confine ourselves to a brief presentation of the two levels that are most relevant to our analysis: the interpersonal and the representational level. For a complete outline of FDG we refer to Hengeveld (2005).

3.2 The interpersonal level

At the interpersonal level all relevant units of communicative behaviour are formalized in terms of their communicative function. The overall structure of this level is given in Figure 1.

\[
(M_1: [(A_1: [(F_1) (P_1)^N (C_1: [(T_1) (R_1)^N \ldots]) (C_1)]) (A_1)]) (M_1))
\]

Figure 1. The interpersonal level

The hierarchically highest unit of analysis given here is the move (M). A move may contain one or more (N) discourse acts (A). A discourse act consists of an illocution (F), one or more speech act participants (P), and the communicated content (C) presented by the speaker. The communicated content, in its turn, may contain a varying number of ascriptive (T) and/or referential (R) acts. Note that all units within a pair of square brackets are operative at the same layer, i.e. there is no hierarchical relation between them. Variables at the interpersonal level are given in capitals for ease of recognition.

3.3 The representational level

The internal, hierarchically layered structure of the representational level is presented in Figure 2.

\[
(p_1: [(e_1: [(f_1) (x_1) (l_1) (t_1)^n \ldots] (e_1)]) (p_1))
\]

Figure 2. The representational level

At this level of analysis linguistic units are described in terms of their semantic category. The highest layer here presented is the propositional content (p). A propositional content is a mental construct, and may contain one or more descriptions of states-of-affairs (e). Within the description of a state of affairs various other semantic categories may enter, including properties (f), individuals (x), spatial regions (l), and temporal

4. Upward layering in units of higher order is possible, but not relevant for the purpose of this paper.
regions (t). Note that this level is purely descriptive in nature, i.e. linguistic units are described in terms of their designation. The use that is made of these units, for instance reference or ascription, is accounted for at the interpersonal level. Again, square brackets set off sets of units operating at the same layer.

3.4 Heads, modifiers, operators, functions

Each of the units discussed so far may be expanded in the following way, where α ranges over all variables:

\[(\pi_1: [(\text{complex head}) (\alpha_1): \sigma (\alpha_1)])_q\]

A unit may be built up using lexical and grammatical means. The lexical means can be subdivided into obligatory heads and optional modifiers (σ). The head is represented as the first restrictor, the modifier as a non-first restrictor. Heads may be complex, as when a number of coordinated units together define a hierarchically higher unit. Simple heads are lexemes occupying the first restrictor slot. Modifiers may again be classified in terms of their semantic category, i.e. they may designate spatial regions (l), temporal regions (t), etc. Grammatical means are subdivided into operators (π) and functions (φ). Operators capture non-relational properties expressed through grammatical means, while functions capture relational properties expressed through grammatical means.

By way of example, consider the interpersonal (37) and representational (38) formalization of example (36):

\[(36) \text{Reportedly a man cut himself with a knife yesterday.}\]

\[(37) (A_I: [(F_I: \text{DECL } (F_I)) (P_I)S (P_I)_A (C_I): [(T_I) (-\text{id } R_I) (R_I) (-\text{id } R_K) (R_L)] (C_I): \text{reportedlyAdv } (C_I)))] (A_I))\]

\[(38) (p_i: [(\text{past } e_i; [(f_i: \text{cut } (f_i)) (1 \ x_i: \text{man}_N (x_i))_{Ag} (x_i)_{Pat} (1 \ x_j: \text{knife}_N (x_j)_{Instr}] (e_i); (t_i; \text{yesterdayTempAdv } (t_i)) (e_i)))) (p_i))\]

At the interpersonal level, the act A_I has a complex head, consisting of a series of hierarchically equivalent units in between square brackets. The illocution F_I has an abstract head representing the basic illocution of the discourse act. The speech act participants are provided with functions indicating their role within the discourse act. The communicated content C_I has a complex head, but also a modifier reportedly indicating that the communicated content was obtained from another speaker. Within the communicated content there is a series of subacts: one ascriptive act, corresponding to cut; and four referential acts, corresponding to a man, himself, a knife, and yesterday, respectively. The first and third of these carry the operator -id ‘non-identifiable’, which triggers the indefinite expression of the noun phrases.

At the representational level, the act A_I has a complex head and no modifiers. Similarly, the state-of-affairs e_i has a complex head, but it is modified by yesterday, which itself designates a temporal region (t_i). Time is also expressed through
the past tense, which is captured by the operator \textit{past}. Within the description of the state of affairs there are several units, one designating a property \((f_i)\) and three designating individuals \((x_i, x_j, x_k)\). The specific roles these individuals play within the state of affairs is indicated by means of the functions \textit{Ag} ‘agent’, \textit{Pat} ‘patient’, and \textit{Instr} instrument.

4. Conjunctions at the representational level

Using the framework just introduced, we now return to the distinction between grammatical and lexical conjunctions. In this section we will first offer a formalization of this distinction at the representational level. In Section 5 we then turn to the use of conjunctions at the interpersonal level.

In developing our proposal we take our lead from Mackenzie’s (1992a, 1992b, 2001) work on adpositions, in which he similarly draws a distinction between lexical and grammatical adpositions, representing the former as lexical heads of (locative) phrases, and the latter as the expression of semantic functions. Applying this to the simple conjunctions discussed in Section 2, we obtain the results given in (39)–(42). Note that we have simplified the representations where possible, leaving out details that are irrelevant to our argumentation.

(39) She called him \textit{before} she left.
(40) \((e_i: [\text{she called him}] (e_i): (t_i: (f_i: \text{before}_\text{conj} (f_i)) (t_i)\cap (t_j: (e_j: [\text{she left}] (e_j) (t_j)_{\text{ref}})) (e_i)))\)
(41) She stayed home \textit{until} the meeting began.
(42) \((e_i: [\text{she stayed home}] (e_i): (t_i: (t_j: (e_j: [\text{the meeting began}] (e_j): (t_i)\text{all} (t_i)))) (e_i)))\)

The dependent clause \textit{before she left} as a whole is an optional addition to the main clause and is therefore represented as a modifier occupying the second restrictor slot. This modifier designates a temporal region, and is therefore provided with the variable \(t_i\). This temporal region is defined in relation to the temporal region \(t_j\) of the event \textit{she left}, which, being a state of affairs, is provided with the variable \(e_j\). The lexical conjunction lexically specifies the relation between the temporal regions \(t_i\) and \(t_j\) and is therefore provided with the variable \(f_i\), which is used for properties and relations. It designates a two-place relation, and the functions of the two units that are in this relation are \textit{O} ‘zero’, for the bearer of a property, and \textit{Ref} ‘reference’, for the entity in relation to which the property is defined.

The dependent clause \textit{until the meeting began} in (41) is again an optional addition to the main clause, and therefore formalized in (42) as a modifier, designating a temporal region \(t_j\). In this case the conjunction is grammatical in nature, and can therefore be represented as a function \textit{All} ‘allative’ of the unit designating another temporal region \(t_j\). This region is defined in terms of the occurrence of the event \textit{the meeting began},
which is represented as \( \epsilon_j \). The basic difference between the two constructions, then, is
that in the latter case the temporal region, itself defined by a state of affairs, operates at
the level of the main clause, whereas in the former case it is embedded as an argument
of the phrase headed by \( \text{before} \).

These representations help us capture the differences in behaviour that were noted
in Section 2. We noted first of all that there are two types of modifiers of adver-
bial clauses, one with narrow scope and one with wide scope. The following example
contains both:

(43) He arrived exactly three hours before she left.

The underlying representation of this example shows the difference in scope between
the modifiers. Note that we use the variable \( q \) here for quantity/measure.

(44) (\( e_i: \text{[he arrived]} \) (\( e_i \)): (\( t_i: (\text{fi: beforeConj} (\( fi: \) three hours (\( qi: \)) (\( fi \))) (\( ti \)):\() \)) (\( ej: \text{[she left]} \) (\( ej \)) (\( tj: (\text{Ref: exactly} (\( ti \))) (\( ej \))\))

In the case of grammatical conjunctions a lexical head is absent, and there is therefore
no slot available for a narrow scope modifier either, which means that the formaliza-
tions proposed here correctly capture the empirical observations.

The second observation that we made earlier concerned the combinability of lexi-
cal and grammatical conjunctions. On the basis of the distinct formalizations of these
two types of conjunctions, example (45) may be represented as in (46):

(45) She stayed until three hours after he left.

(46) (\( e_i: \text{[she stayed]} \) (\( e_i \)): (\( t_i: (\text{ti: afterConj} (\( fi: \) three hours (\( qi: \)) (\( fi \))) (\( ti \)):\() \)) (\( e_j: \text{[she left]} \) (\( ej \)) (\( tk: (\text{Ref: all} (\( ti \))) (\( ej \))\))

The representation of the time region \( t_j \), formalizing the part \( \text{three hours after she left} \),
is provided here with the allative semantic function, which is expressed as \( \text{until} \). Since
\( \text{after} \) is not the expression of a function but a lexical element, the function slot is avail-
able for grammatical markers, which correctly reflects the fact that the two may be
combined in (45).

We now turn to complex lexical conjunctions. We do not discuss these here in
contrast with complex grammatical conjunctions, as the latter would be just as much
the expression of functions as simple grammatical conjunctions. Complex lexical con-
junctions come in various types, illustrated in (47)–(48):

(47) The moment (that) he arrived in London it started raining.

(48) Smallpox would be rapidly controlled \( \text{in the event} \) that it were introduced into
Australia.

The difference between the two constructions is that in (47) the \( \text{that} \)-clause modifies,
i.e. further specifies, the head \text{moment} in terms of an event taking place. It is therefore
similar to a relative clause. This is not true of (48), in which the \( \text{that} \)-clause refers to
the same event as the head event, although in a more specific way. It is therefore similar to a restrictive apposition (see Mackenzie 1990 for a discussion of this distinction).

Given these characteristics, (47) may be represented as follows:

\[(49) (e_i: \text{[it started raining]} (e_j: (t_i: (f_i: \text{moment} (f_i)) (t_i): (e_j: \text{[he arrived in London]} (e_i): (t_i) (e_i))))\]

What this representation captures is that this is a case of relativization on the temporal specification of the event contained in the relative clause modifying \text{moment}. The temporal adverbial may then be paraphrased as 'the moment such that \text{ej} took place at that moment'. Note that \text{moment} is treated as a noun and not as a specialized conjunction.

Given its restrictive appositional nature, (48) may be represented as:

\[(50) (e_i: \text{[small pox would be rapidly controlled]} (e_j: (f_i: \text{event} (f_i)) (e_j): (e_k: \text{[smallpox are introduced into Australia]} (e_k)) (e_j)) (e_i))\]

In this case the event description \text{ej} that has event as its head is further specified by another event description \text{ek}, which gives the more specific description. The preposition \text{in} is triggered by the locative semantic function, which here has to be interpreted metaphorically as it is applied within the domain of event descriptions.5

A major difference between (47) and (48) is that in (47) the clause modifying the head noun \text{moment} is a predication with an open time slot in the modifier position, i.e. \text{moment} plays a role within the embedded clause itself, whereas in (48) the modifying clause is a closed predication, i.e. \text{event} plays no role within the embedded clause.

A last case that merits some attention is the following:

\[(51) \text{The facilities can be used immediately that you join.}\]

This case is quite similar to the previous one, in the sense that both \text{immediately} and \text{you join} are alternative descriptions of the same time span, the latter one being more specific than the former. We may therefore formalize this example as another case of apposition, though of units designating temporal regions rather than states of affairs:

\[(52) (e_i: \text{[the facilities can be used]} (e_j: (t_i: (f_i: \text{immediately} (f_i)) (t_i): (t_j: (ej: \text{[you join]} (e_j)) (t_j) (t_i)) (e_i)))\]

Generalizing over the various cases of lexical conjunctions and periphrastic conjunctional phrases that have been discussed in this section, we find that next to a limited number of lexical items that are specialized in conjoining and therefore have to be identified as belonging to a lexical class of conjunctions, complex lexical conjunctions exploit existing lexical categories of the language in order to indirectly express a semantic relation between clauses. In all cases discussed grammatization of the construction leads to a situation in which the internal complexity of the construction

5. Note that \text{in} is a lexical preposition in its basic locative use, but a grammatical one in its metaphorical use.
is reduced and the semantic relation between clauses is established directly through a grammatical element.

5. Conjunctions at the interpersonal level

In the previous section we discussed various conjunctions and conjunctional phrases operating at the representational level. Conjunctions may also, however, fulfil a role at the interpersonal level. Consider the following examples:

(53) Watch out, because there is a bull in the field.
(54) He is a nice guy, although you probably knew that already.

What the adverbial clauses in (53)–(54) have in common is that they in one way or another comment on the appropriateness of the discourse act expressed in the main clause. They thus qualify an interpersonal rather than a representational unit. One way to show this is that in these uses they can never be in the scope of a semantic modifier, not even the ones with wide scope discussed in 2.4:

(55) *Watch out, exactly because there is a bull in the field.
(56) *He is a nice guy, exactly although you probably knew that already.

This difference becomes even more evident in those cases in which a conjunction may be used at both the representational and the interpersonal level. Compare the following example with (55):

(57) Providing food assistance is not easy exactly because the infrastructure is lacking.

In the strictly causal use in (57) the phrasal modifier exactly is allowed, whereas in the argumentative use in (55) it is not. These facts can be seen as a reflection of the fact that in its causal use because operates at the representational level, and can thus be modified by semantic phrasal modifiers, whereas at the interpersonal level it is outside the scope of semantic modification. The difference between the two uses may be formalized as in (58) and (59):

(58) (e_i: [providing food assistance is not easy] (e_i): [the infrastructure is lacking (e_i)Cause: exactly (e_i))
(59) (M_i: (A_i: [watch out] (A_i)), (A_j: [there is a bull in the field] (A_j))Motivation) (M_i))

The causal clause in (57) is formalized in (58) as a modifier, since it is used restrictively. It is provided with the semantic function Cause, which triggers the causal grammatical conjunction because. The argumentative clause in (59), which represents (53), is paratactically related to the main clause, and carries the rhetorical function Motivation, reflecting the argumentative use of the grammatical conjunction because. It is
represented as a separate discourse act. The act status of this clause is evident from the fact that it allows for interpersonal modification with scope over the *because* clause as a whole, as illustrated in (60):

(60) Watch out, because, **frankly**, there is a bull in the field.

The act status of the *because*-clause in (53) is furthermore reflected in the fact that it is not and cannot be within the scope of the basic illocution of the main clause, and there is even a (limited) possibility of changing the basic illocution of this clause, as illustrated in (61):

(61) Watch out, because aren’t bulls dangerous?

The possibility of modifying the basic illocution of the *because* clause reveals the presence within this clause of an illocutionary component, which is a defining characteristic of discourse acts.

A remarkable fact about adverbial conjunctions that are employed at the interpersonal level is that they are highly grammaticalized, as reflected in the impossibility of modification (see 2.4). Typical examples of conjunctions frequently used at the interpersonal level in English are *although, since, for, and so*. This contrasts sharply with the wide range of lexical conjunctions and conjunctural phrases operating at the representational level, several of which were discussed in the previous section. This high degree of grammaticalization of interpersonal conjunctions is not unexpected. It is a general prediction in FDG that grammatical elements, when they have their origin at the inner layers of the hierarchical structure of the clause, develop increasingly abstract functions at next higher layers and levels of linguistic organization (see e.g. Hengeveld 1989).

A group of seeming counterexamples to this generalization is exemplified in (62) and (63):

(62) **Considering** that salaries are rising, we should try to reduce the production costs.

(63) Keep your money (**assuming** that you have any) separate from mine.

The phrases *considering that* and *assuming that* introduce argumentative steps in the discourse, representing pieces of background knowledge and background assumptions that lead to certain conclusions. Other such phrases are *supposing that, granting that, given that, and provided that*. Within these phrases seemingly functioning as conjunctural elements the verbal form can be modified in a limited number of ways:

(64) **Considering further** that salaries are rising, we should try to reduce the production costs.

(65) Keep your money (**assuming for a moment** that you have any) separate from mine.
This lexical behaviour of these conjoining expressions functioning at the interpersonal level seems to contradict our earlier generalization concerning the grammatical status of conjunctions at the interpersonal level.

In our analysis, however, the above constructions are not interpreted as conjunctive phrases. Recall that in our discussion of example (3) we argued that participial forms can be used in narrative chaining. The conjoining expressions we are discussing now take a participial form as well, but they are not used in narrative chaining but rather in argumentative chaining. Within this chaining construction the verbs involved remain lexical, acting as the head of cosubordinate clauses, within which they can be modified. The dependent status of the cosubordinate clauses does not reside in the verb as such, but in the participial ending with which the verb is provided. The difference with the cases of narrative chaining discussed earlier is that the chained units are not states of affairs (e) at the representational level but discourse acts (A) at the interpersonal level. The difference may be represented schematically as in (66) and (67):

(66) \((e_1), \ldots, (e_n)\)

(67) \((A_1), \ldots, (A_N)\)

A major reason to assume the underlying representation in (67) is that a dependent clause such as the one in (62) cannot be within the scope of the basic illocution of the main clause, as illustrated in (68):

(68) Considering that salaries are rising, shouldn’t we try to reduce the production costs?

On the other hand, it does not seem to be possible to modify the illocution of the dependent clause itself:

(69) *Considering that aren’t salaries rising, we should try to reduce the production costs.

Our solution for this asymmetrical behaviour of the two acts involved is to assume that in (62) and (68) consider itself occupies the illocutionary slot, just like performative verbs do, for consider describes a presentative/discourse property of the content of the clause. The underlying representation of (68) would then be as in (70), where the function Prep ‘preparation’ marks the preparatory act and the function Nucl ‘nuclear’ the nuclear act:

(70) \((A_i: ([F_1: considerV (F_1)) (P_1)(S (C_1))] (A_i))_{Prep}, (A_j: ([F_1: INT (F_1)) (P_1)(S (P)_A (C_j))] (A_j))_{Nucl}\)

Every C is then mapped onto a propositional content at the representational level.

This analysis not only solves the problem of the observed asymmetry in the behaviour of the two discourse acts, it also explains why only certain types of modifier can be used with consider when used in this function. Consider the following examples:
(71) Considering for a moment/further/for the sake of the argument that salaries are rising, we should try to reduce the production costs.

(72) *Considering for the last three hours/for Bill’s sake/in the director’s office that salaries are rising, we should try to reduce the production costs.

The only modifications allowed, as illustrated in (71), are directly related to the speech situation itself, rather than to the features external to that speech situation specified in the ungrammatical constructions in (72). This is a direct reflection of the interpersonal status of consider in this configuration of acts.

The regular association between an interpersonal argumentative function and a limited set of argumentative verbs may be expected to induce a process of grammaticalization which leads to the reinterpretation of the participial verb forms involved into interpersonal conjunctions. There are certain indications that this process is actually taking place, since, as noted by Quirk et al. (1985:1002f.), some of the participial forms used in the constructions discussed here ‘retain certain properties characteristic of verbs, while those that are most like simple conjunctions have lost all such properties’. They show that some participial forms do allow adverbial expansion, whereas others do not:

(73) Supposing/assuming for the sake of the argument that . . .

(74) *Seeing/provided for the sake of the argument that . . .

Once a participial form grammaticalizes into a conjunction, it no longer forms part of a chaining construction, but of a subordinating construction, and the formalization in (70) changes into for instance the one given in (75):

(75) (M₁; [(A₁) (A₂)Assumption] (M₂))

Thus, whereas the process of grammaticalization at the representational level seems to draw mainly on a specific group of nouns, the process of grammaticalization at the interpersonal level seems to draw mainly on a specific group of verbs. In the latter case, however, the verbs involved start out in their regular function at the representational level, and only move up to the interpersonal level once they grammaticalize into conjunctions.

6. Conclusions

In this paper we have tried to argue two main points with respect to the analysis of English conjunctions within Functional Discourse Grammar. First of all, we argued that within the class of conjunctions a distinction has to be made between grammatical and lexical conjunctions, parallel to the distinction between grammatical and lexical prepositions proposed in Mackenzie (1992a, 1992b) and later publications. Secondly, we argued that lexical conjunctions play a role at the representational level only. The analysis of seeming counterexamples involving the use of participial verb forms has
led us to introduce the concept of ‘argumentative chaining’ next to ‘narrative chaining’, which provides a parallel analysis of chaining constructions at the representational and interpersonal level within Functional Discourse Grammar.

References


Tree tigers and tree elephants

A constructional account of English nominal compounds*

Matthew Anstey
Charles Sturt University, Canberra

The description of English nominal compounds (ENCs) poses a challenge to linguistic theories because ENCs compress a bewildering array of semantic and conceptual information into a relatively simple syntactic structure. The present contribution reviews linguistic and psycholinguistic efforts to meet this challenge in a variety of ways. It is proposed that a superior explanation is one that treats the semantic and syntactic (and conceptual and prosodic) information as a constructional unity, learned and used as a distinct linguistic entity rather than as a derivative of construction-independent rules and processes. The proposed ENC construction is claimed to have theoretical, functional, and cognitive adequacy, a desideratum of linguistic theories.

* It is a privilege to contribute the present article in a book to honour Lachlan Mackenzie. The oft-quoted line from Albert Camus’s *The Rebel*, 'Real generosity towards the future lies in giving all to the present,' sums up for me Lachlan’s approach to life, as I have found him to be singularly generous with his time, resources, learning, friendship, and editorial meticulousness.

I had planned to include some constructional reflections on Functional Discourse Grammar (FDG), a theory of interest to us both, so although it was not to be, I hope this contribution can nevertheless contribute to FDG’s development.

I would like to thank the following people for commenting on earlier drafts: Christina Gagné, Edward Wisniewski, William Croft, Malcolm Ross, the editors, and an anonymous reviewer. I also wish to thank the United Bible Societies for their financial assistance during my doctoral studies. Correspondence can be addressed to manstey@csu.edu.au.
1. Introduction

*tree claim*, n. (1890, US English): “a piece of land allotted with the proviso that it shall become the property of the occupier after a fixed term on condition of his planting a certain proportion of it with trees” *Oxford English Dictionary*

English nominal compounds (ENCs) such as *tree claim* illustrate why these innocuous constructions have received so much attention: their syntactic simplicity belies a seemingly limitless inventory of meanings, generated not only by semantic relationships between the component nouns but also by unpredictable conceptual blends. In terms of English clause structure and function, ENCs – clauses in miniature – remind us of Vendler’s description of nominalization as “packing a sentence into a bundle that fits into other sentences” (1967, quoted in Lynott & Keane 2003).

Moreover, the complexity of these bundles works at two levels: first, there is the taxonomic complexity (see Section 2) of ENCs as a collection, with new candidates regularly vying for semantic novelty (e.g. *hockey dads, the Blair camp*; cf. Moder 2004). Moreover, subtle constraints stifle excessive creativity (e.g. *tree tiger* vs. *tree elephant, traffic congestion* vs. *traffic indigestion*). Secondly, for many ENCs there are often several interpretations (e.g. *chocolate skin, cactus friends*). How does context work to assist the correct interpretation?

In addition to the onerous task of adequately describing such data, one must account for the empirical observation that the average English speaker manages with ease to generate and interpret ENCs, somehow navigating through this complexity with remarkable speed and success. Thus, if functional theories of language are to reach their self-declared goals of psychological adequacy (Dik 1997; Van Valin & LaPolla 1997; Croft 2001), explaining the ENC phenomena is a desideratum.

By considering the psycholinguistic research on ENC comprehension, the present article approaches the problem of ENCs from the perspective of the listener/reader. It begins with an overview of ENCs, illustrating their semantic diversity and briefly profiling three theories of ENCs (Section 2). This is followed by a review of psycholinguistic research on ENC comprehension (Section 3). Section 4 presents a model of ENC comprehension (and production) that accounts for the linguistic and psycholinguistic data presented in Sections 2–3, a model influenced by Construction Grammar (Croft 2001, 2005; Goldberg 1995, 2003; Fried & Östman 2005).

2. An overview of ENCs

These compounds are a microcosm of the generative nature of natural language, in which we see new meanings being created from the re-combination of words in syntactically well-formed phrases. It is the need to understand this generativity of language that has motivated several decades of research into nominal compounds. (Lynott & Keane 2003)
The syntax of ENCs at first blush is simplicity itself: ostensibly they have a syntactic structure of \([N N]_X\), in which the second noun is the head, regardless of one's definition of headedness. Giegerich (2004; also Libben et al. 2003 for semantic transparency), nevertheless, notes many subtle differences between phrasal (steel bridge) and lexical ENCs (silverfish), as indicated in Table 1.

Haskell et al. (2003) also review a further semantic peculiarity of phrasal ENCs—the modifying noun can only be an irregularly formed plural (mouse trap, mice trap, rat trap, *rats trap), except in cases of semantically multiple-type abstract modifiers (awards ceremony, pilots union, sports announcer, weapons inspector). It is the use of the plural to indicate multiple types rather than multiple tokens that is crucial here (i.e. an awards ceremony bestows a variety of awards but an award ceremony bestows a particular award to many people). Libben et al. (2003) also note that type interpretations only work for phrasal ENCs (a steel bridge is a type of bridge vs. *silverfish is a type of fish).

Obviously, many of these distinctions stem from the headedness of [steel [bridge]]Head versus [silverfish]Head.

Thus phrasal ENCs, the topic of the present paper, might surface with a simple structure, but the full range of their semantic, syntactic, and prosodic features is by no means simple.

The semantics of ENCs, moreover, are famously diverse. Consider a random sample of ENCs, most of which are or at least were once phrasal, modified by tree, tiger, and elephant—the three nouns in the article's heading—from the Oxford English Dictionary (Table 2).

Wisniewski and Middleton (2002: 1; cf. Wisniewski 1996; Costello & Keane 2001) identify three basic semantic types: relation compounds “involve a thematic relation between the referents of the modifier and head nouns”, so robin hawk (also, apartment dog) is interpreted as ‘a hawk that preys on robins’; property compounds involve the application of “one or more properties of the modifier noun to the head noun,” so robin hawk (also, elephant fish) is interpreted as ‘a hawk with a dull red breast’; and

---

1. Giegerich’s (2004) article on stress investigates the correlations between fore-/end-stress and lexical/phrasal ENCs respectively.
Table 2. ENCs with *tree*, *tiger*, and *elephant* modifiers

<table>
<thead>
<tr>
<th>Compound</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>tree avenue</td>
<td>‘an avenue of trees’</td>
</tr>
<tr>
<td>tree dweller</td>
<td>‘one who dwells in trees’</td>
</tr>
<tr>
<td>tree cactus</td>
<td>‘a tall cactus, as the saguaro’</td>
</tr>
<tr>
<td>tree orchid</td>
<td>‘an orchid growing on trees’</td>
</tr>
<tr>
<td>tree tiger</td>
<td>‘leopard’</td>
</tr>
<tr>
<td>tree butterfly</td>
<td>‘a butterfly that lives among trees’</td>
</tr>
<tr>
<td>tree agate</td>
<td>‘an agate with dendritic markings’</td>
</tr>
<tr>
<td>tree claim</td>
<td>(1890) ‘a piece of land allotted with the proviso that it shall become the property of the occupier after a fixed term on condition of his planting a certain proportion of it with trees’</td>
</tr>
<tr>
<td>tree god</td>
<td>‘a divinity that is supposed to inhabit a tree’</td>
</tr>
<tr>
<td>tree line</td>
<td>‘the line on a mountain above which trees do not grow’</td>
</tr>
<tr>
<td>tree lore</td>
<td>‘lore about trees’</td>
</tr>
<tr>
<td>tree planter</td>
<td>‘one who plants trees’</td>
</tr>
<tr>
<td>tree lily</td>
<td>‘a tree with lily-like flowers’</td>
</tr>
<tr>
<td>tree monkey</td>
<td>‘a monkey that lives in a tree’</td>
</tr>
<tr>
<td>tree crab</td>
<td>‘a crab that lives in a tree’</td>
</tr>
<tr>
<td>tree bug</td>
<td>‘an insect that feeds upon tree juices’</td>
</tr>
<tr>
<td>tree box</td>
<td>‘a frame used to protect small trees’</td>
</tr>
<tr>
<td>tree worship</td>
<td>‘the worship of trees’</td>
</tr>
<tr>
<td>tree house</td>
<td>‘a house built in a tree’</td>
</tr>
<tr>
<td>tree surgery</td>
<td>‘the pruning, repair, and preservative treatment of trees’</td>
</tr>
<tr>
<td>tiger trap</td>
<td>‘a trap that traps a tiger’</td>
</tr>
<tr>
<td>tiger beetle</td>
<td>‘striped beetle’</td>
</tr>
<tr>
<td>tiger suit</td>
<td>‘a striped combat uniform worn as camouflage’</td>
</tr>
<tr>
<td>tiger cage</td>
<td>‘a cage designed to contain a tiger’</td>
</tr>
<tr>
<td>tiger fury</td>
<td>‘a ferocious fury’</td>
</tr>
<tr>
<td>tiger shark</td>
<td>‘a voracious shark’</td>
</tr>
<tr>
<td>elephant bird</td>
<td>‘large fossil bird’</td>
</tr>
<tr>
<td>elephant joke</td>
<td>‘a child’s nonsense joke of which an elephant (usu. in a ridiculous situation) is the subject’</td>
</tr>
<tr>
<td>elephant grass</td>
<td>‘a type of grass’</td>
</tr>
<tr>
<td>elephant beetle</td>
<td>‘large South American beetle’</td>
</tr>
<tr>
<td>elephant-gravel</td>
<td>‘gravel containing remains of elephants’</td>
</tr>
<tr>
<td>elephant-path</td>
<td>‘a path trodden by elephants’</td>
</tr>
</tbody>
</table>

Hybrid compounds (conjunctive compounds in Costello & Keane 2001) refer to a mixture of two entities, so *robin hawk* (also *pet bird*, *singer songwriter*) is interpreted as ‘a bird that is a cross between a hawk and a robin’.

In addition to syntax and semantics, pragmatics also has a critical role in interpreting ENCs, in the sense that the context provides interpretative clues (Moder 2004). As phrasal ENCs drift to lexical ENCs, abstruse contextual information can become reified into the meaning, so that not even a generic \( N = a \text{ that } V(PP) \ N \) will provide...
a parsing strategy for the listener. Some such examples from Table 2 are tree claim, tiger suit, and elephant joke. N wine compounds are similarly complex, for example, ice wine, is defined in OED (under Eiswein, n.) as “wine made from ripe grapes picked while the frost is on them, and still frozen when they go into the press” and dessert wine is something like “wine having a sweetness typically associated with sweet desserts, typically drunk with dessert”.2

Ryder (1994:8–11) divides ENCs in a different way, according to usage: deictic, novel, and established compounds. Deictic compounds are “created to satisfy a fleeting discourse need”. The interpretation of these is normally entirely dependent on the context providing the necessary background information. Ryder gives the example of a group of apartment dwellers at a meeting to convince management to bar certain pets from their complex. Each group member has been chosen to represent a particular pet. In this context one could say, Maggie is our cat person. What makes cat person unusual is that it is extremely difficult without a context to generate an N who does not like N compounds.

Corpus examples of deictic ENCs are easy to locate:

(1) She watched while her friend began what they both called her tiger prowl, that compulsive pacing up and down, head lowered, hands sunk in her dressing-gown pockets. (BNC G3E 01096 = Death of an expert witness. James, P. D., Sphere Books Ltd, London [1979]).

Novel compounds are used typically among a restricted group of people, often as naming devices, as in Downing’s example of bike girl meaning ‘the girl who parked her bike in the vestibule’ (Downing 1975 quoted in Ryder 1994:9). The interpretation of novel compounds has preoccupied psycholinguistic research into ENCs. The third type is established compounds, such as horseradish, which are lexicalized to various degrees.

Undoubtedly, the most remarkable pragmatic-semantic feature of ENCs is their ability to compress much information into a simple structure. A striking example of this, peculiar to Australian/New Zealand English, is ENCs of the form Claytons N, meaning ‘an N you have when you are not having an N’. Claytons™ is a non-alcoholic beverage that ran a famous television advertisement (from the late 1970s) with the slogan, “Claytons: The drink you have when you’re not having a drink”. This proper name plus noun compound is very productive. A quick search of “Claytons” on Google™ provided many examples, the first two in Table 3 demonstrating the “compositional” unpacking (which is of course non-compositional).

The last two examples, respectively, are from a government media release and a Senate submission to the Australian government on terrorism insurance, demonstrating how acceptable this nominal modifier has become. (An example was also found

2. Even the simple white wine is not straightforward (i.e., it is not a case of x is wine and x is white). The modifier has unexpectedly lost its semantic transparency, which might explain why many people incorrectly think white wine and red wine are made from white and red grapes respectively.
Table 3. Claytons N Australian/New Zealand ENCs [emphasis MPA]

<table>
<thead>
<tr>
<th>Example</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claytons Marketing … The Marketing …that you have when you do not want a full time position.</td>
<td><a href="http://www.colourthinking.com.au/">http://www.colourthinking.com.au/</a> claytons_marketing_and.htm</td>
</tr>
<tr>
<td>This is a mystery only God can work out. It is what I term a Claytons pregnancy (a pregnancy that you are having and the doctors say you are not).</td>
<td><a href="http://www.pregnancy.com.au/">http://www.pregnancy.com.au/</a> clayton's_pregnancy.htm</td>
</tr>
<tr>
<td>Claytons Convicts (for those who don’t own one).</td>
<td><a href="http://members.iinet.net.au/~perthdps/convicts/diycons.html">http://members.iinet.net.au/~perthdps/convicts/diycons.html</a></td>
</tr>
<tr>
<td>Claytons Fried Fish … The fried fish that isn’t fried!</td>
<td><a href="http://www.abc.net.au/northwest/stories/s1076987.htm">http://www.abc.net.au/northwest/stories/s1076987.htm</a></td>
</tr>
<tr>
<td>Consumers will be given a Claytons choice between herbs and spices laced with a potential cancer-causing agent or zapped with nuclear [sic] if the Australia New Zealand Food Authority has its way this week.</td>
<td></td>
</tr>
<tr>
<td>The US, France, Spain and Germany have become reinsurers of last resort against terrorist activity over the past eighteen months. Many other European countries are moving toward intervention frameworks. They recognise that the Claytons cover on offer provides totally inadequate protection (see initial Property Council Submission).</td>
<td><a href="http://www.aph.gov.au/Senate/committee/economics_ctte/terrorism/submissions/sub06a.pdf">http://www.aph.gov.au/Senate/committee/economics_ctte/terrorism/submissions/sub06a.pdf</a></td>
</tr>
</tbody>
</table>

in an economics PhD dissertation). The reason is obvious: Claytons N packs a lot of information into a single ENC. And yet, it is still more phrasal than lexical (a Claytons pregnancy and a real one; a [Claytons [paper box]]; [Claytons X] is productive, and so forth).

A search of BNC revealed additional examples of pragmatic compression, such as in the following examples, which could hardly be argued to be lexical:

1. Ciccolini is not a brainstorming piano tiger, though his playing is never less than exciting and this series makes an excellent bargain, particularly as the competition is far from extensive. (BNC BMC 01218 = CD Review. CD Review Ltd, Berkhamstead [1992–1995]).
Tree tigers and tree elephants

(3) True, they do not get any cash, but they get a big chunk of valuable stock and, most important, they get to leave the Unix business and stop supporting the big white Unix elephant. (BNC CSS 00342 = Unigram X. APT Data Services Ltd. [1993-04/1993-05]).

Piano tiger can be construed as 'a tiger who plays piano' where 'tiger' is metaphor for 'person with prowess', resulting in 'a person skilful at playing piano'. The big white Unix elephant is a curious combination of the lexicalized compound white elephant modified internally by Unix, a type of computer software, which together conveys 'the big "white elephant" that is the operating system Unix' where 'white elephant' stands for 'cumbersome, unprofitable object', resulting in 'the large, cumbersome, unprofitable Unix operating system.'

2.1 Linguistic theories of ENCs

There is not the space to review the large number of linguistic descriptions of ENCs (see Ryder 1994:19ff.). They are, in my view, variations on a theme of solutions to the missing information problem; that is, ENCs compress something into a unit to produce a new designation, the something being broadly understood as information: What is this missing semantic, syntactic, pragmatic information? Where (and/or when) is it deleted? How is it recovered? How do speakers and addressees rapidly and (on the whole) successfully move between complex conceptualizations and ENCs?

Three examples illustrate the preceding claim. First, Levi (1978) proposed that all ENCs could be derived from underlying syntactic structures (USs) with one of nine Recoverably Deletable Predicates (RDPs): cause, have, make, use, be, in, for, from, and about. Furthermore, the modifier/first noun is either the direct object or subject of the relative clause in the US, as illustrated in Table 4.

<table>
<thead>
<tr>
<th>RDP</th>
<th>modifier as direct object</th>
<th>modifier as subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>cause</td>
<td>tear gas</td>
<td>nicotine fit</td>
</tr>
<tr>
<td>have</td>
<td>picture book</td>
<td>government land</td>
</tr>
<tr>
<td>make</td>
<td>silkworm</td>
<td>daisy chain</td>
</tr>
<tr>
<td>use</td>
<td>steam iron</td>
<td></td>
</tr>
<tr>
<td>be</td>
<td>soldier ant</td>
<td></td>
</tr>
<tr>
<td>in</td>
<td>field mouse</td>
<td></td>
</tr>
<tr>
<td>for</td>
<td>horse doctor</td>
<td></td>
</tr>
<tr>
<td>from</td>
<td>olive oil</td>
<td></td>
</tr>
<tr>
<td>about</td>
<td>tax law</td>
<td></td>
</tr>
</tbody>
</table>

3. Janssen (this volume) similarly identifies the “open-ended interpretability and the interpretive dependence on the speaker and hearer” of novel nominal compounds as two of the central issues to be explained in any functional-cognitive theory of such.
For example, thermal stress is derived from [stress [CAUSE heat stress]] (Levi 1978: 121–123). Levi clearly places the bulk of the missing information in syntax: this motivates the restriction of underlying subjects to only CAUSE, HAVE, and MAKE (see Table 4), because these three RDPs are the only predications allowing passivization, a crucial transformational stage in her derivational model.

Secondly, despite claims from some (e.g. Downing 1977) that the semantic taxonomy of ENCs is without limit, others such as Warren (1978) pursue a semantic inventory approach, trying to derive all ENCS from a small set of basic relations, such as whole-part (armchair), place-object (sea port), copula (girlfriend), and so forth. Our second illustration of the missing information problem, accordingly, is Fabre and Sébillot (1995)'s proposal to characterize ENCs using lexical semantics. Although they incorporate some morphosyntactic information in their model, pride of place goes to a semantic representation projected from an enriched Generative Lexicon (Pustejovsky 1995) enriched with WordNet data (Miller et al. 1990). They explain the need for this as follows (Fabre & Sébillot 1995: 9):

A restricted set of features [in traditional accounts, MPA] proves to be of no help to handle the semantics of noun-noun association, because … the semantic class of the constituents is the only clue to choose between several available predicates. Second, a rich semantic classification is needed to factorize predicative information.

Their interpretative system consists of a linking mechanism between semantico-syntactic constructions and the lexicon. For example, consider the ENC earthquake coverage, with two meanings of ‘insurance coverage for earthquakes’ and ‘reportage about earthquakes’. The ENC is linked to these two construals as follows:

– Coverage is an action deverbal, thus in their notation its morphosyntactic construction is <N> + <V-act>.
– <N> + <V-act> is linked to all lexical entries of the form lemmaV (lemmaN, Role), which has in this case provides two possibilities: cover (lemmaN, Role) and lemmaV (coverage).
– cover (lemma)Theme is listed (in WordNet) under insurance possession; about (coverageN, Communication) lemmaN, Theme) under reportage communication.
– This results in earthquake coverage being linked to two semantic structures, cover (earthquakeTheme) and about (coverageN, Communication, earthquakeN, Theme).

Fabre and Sébillot’s results are impressive at one level: their method successfully interprets many ENCs, including several multi-sense ENCS, such as weekend slayings, wood checker (three senses), backup weapon (two senses), bullet casings, liquor bottles (two senses, one of which is incorrect: ‘a bottle made of liquor’), desk height, and wood ashes. This demonstrates the advances that an enriched lexicon can make; essentially, their system has more information in it, so provided associations can be made between modifier and head, it explains more data. But on another level, their results are discouraging, considering the range of ENCs their system fails to interpret correctly – loss
ratio, weapon workers, art collection, juice exports – due to polysemy, semantic shifts, type coercion, or insufficient decompositional information.

Thirdly, Ryder’s (1994) model is of interest, as it is based on Langacker’s Cognitive Grammar and has been empirically tested through several psycholinguistic experiments. It is understandably more sophisticated, having been developed in a book-length treatment of ENCs. Ryder (1994:195–199, passim) proposes “linguistic templates,” which contain semantic primitives linked to “general schema categories,” which in turn are basic sentence kernels expressing the semantic relationship between the terms. Ryder (1994:189) gives the following example of such a template: “Clothing + Natural Element = Clothing for use when exposed to Natural Element, as in sunglasses, raincoat, snowsuit…” Ryder lists (1994:284) almost fifty general schema categories. For example, participants in one experiment associated tiger man with the following general schema categories (1994:346):

(4) TEND ‘x tends/raises/trains y’
TENDBY ‘x is tended by y’
HUNT ‘x hunts/collects (live animals) y’
HAS ‘x owns y’
LIK ‘x is similar to y’
LIKPT ‘a part of x is similar to a part of y’
LIKPR ‘x is similar to y in personality or mental abilities’
LIKAC ‘x is similar to y in actions’

Ryder also gives numerous examples of a single form having multiple meanings. For instance, [N_{ANIMAL} N_{ANIMAL}] is associated predominately with two templates: [N_{ANIMAL} is similar to N_{ANIMAL}] and [N_{ANIMAL} hunts/collects N_{ANIMAL}] (1994:117). (Note that this example suggests that Ryder treats property compounds as resemblance relation compounds, which, as mentioned below, might be better understood as a separate category.)

Finally, in anticipation of the proposal of Section 4, consider that all three theories can be construed as working with empty constructions containing slots to be or saturated by a variety of syntactic, semantic, and lexical information on their way to expression.

But since such constructions are ubiquitous in grammar, what is so special about ENCs? Absolutely nothing is special about ENCs qua constructions. Rather, it is their high degree of information compression that allures researchers of all persuasions, since ENCs test how well our theories integrate the whole gamut of linguistic knowl-

Table 5. Constructual construal of Levi, Fabre and Sébillot, and Ryder

<table>
<thead>
<tr>
<th>Theorist</th>
<th>Empty</th>
<th>Saturated</th>
<th>Expressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levi</td>
<td>N CAUSE N</td>
<td>stress CAUSE heat stress</td>
<td>‘thermal stress’</td>
</tr>
<tr>
<td>Fabre and Sébillot</td>
<td>lemmaN (lemmaN-Role)</td>
<td>cover (earthquake_theme)</td>
<td>‘earthquake coverage’</td>
</tr>
<tr>
<td>Ryder</td>
<td>x HUNT y</td>
<td>tiger HUNT man</td>
<td>‘tiger man’</td>
</tr>
</tbody>
</table>
edge (cognitive, lexical, semantic, prosodic, etc.) in a psychologically plausible model (albeit only for those committed to psychological plausibility).

For example, Haskell et al. (2003) argue that the previously mentioned semantically-motivated prohibition against morphologically regular modifiers (*rats trap vs. mice trap; but awards ceremony vs. award ceremony), a prohibition evidenced in even 3- to 5-year-olds, poses substantial problems for algorithmic rules-and-exceptions grammars (typically found in generative accounts such as Pinker 1999). Algorithms are by necessity ordered and give rise to bleeding and opacity effects, as is well known in staged models of phonology. So no matter how the rules for ENCs are written, the algorithm will either bleed the possibility for plural modifiers (incorrectly blocking awards ceremony) or not (incorrectly allowing *rats trap). The crucial finding of Haskell et al. (2003:150–154) is however that there is a bias in ENC formation against all plural modifiers, that is, the preference for modifiers is singular > irregular plural > regular plural. But level-ordering accounts like Pinker’s predict that singulars and irregular plurals should act alike because both are stored in the lexicon. Haskell et al. (124) write, “In recognition of some of these limitations [*rats trap vs. awards ceremony – MPA], Pinker (1999) recommended that level-ordering not be taken literally, but rather ‘as laying out the logic of word formation (p. 181)’”, a view that seems to undermine somewhat the internal coherence of the rules-and-exceptions architecture. It is such evidence that further motivates the constructional account of ENCs given in Section 4. That is, the [N-pl N] construction can have specific inherent semantic-conceptual properties, learned by usage.

It is therefore appropriate to turn to psycholinguistic research into ENCs for further illumination.

3. Psycholinguistic accounts

This section reviews the three major psycholinguistic theories of noun-noun interpretation: the Dual-Process model, the CARIN model, and the C³ model. Psycholinguistic models are the outcome of empirical testing of the production and interpretation of nominal compounds and thus offer pertinent insights into the problem.

3.1 The Dual-Process model

Edward Wisniewski and colleagues (Wisniewski 1996, 1997, 1999; Wisniewski & Middleton 2002) have developed the dual-process theory of ENC comprehension,

4. Many functionalist theories, such as Functional Grammar (Dik 1997), Functional Discourse Grammar (Hengeveld 2004), and Role and Reference Grammar (Van Valin & LaPolla 1997), also have algorithmic, or better, modular architectures. Section 4 promotes in contrast a constructional, usage-based architecture.
influenced largely by Murphy (1988, 1990). The Dual-Process model proposes two separate mechanisms, one for generating relational interpretations (a tree snake is a snake that lives in a tree) and another for property interpretations (a tiger clam is a striped clam). The processes run simultaneously (but see Note 15) until a suitable interpretation is found. Estes (2003:315), in separate experiments, confirms the Dual-Process model: “There appear to be distinct attributive and relational processes, and these two processes appear to operate concurrently”. For any given combination, these attributive and relational processes are attempted in parallel and it will be interpreted either attributively or relationally, depending on which process produces a plausible interpretation first.

Wisniewski’s work has focused particularly on the interpretation of property compounds. Wisniewski posits two crucial processes, selection and integration. Selection is the process of choosing a property of the modifier to incorporate into the head noun in the integration process. Like the lexical theories and the other two major psycholinguistic theories, the dual-process model contains an enriched lexicon that allows correct selection. Thus truck has parts wheels, engine, body and events transport, cargo-transport, and so forth. Each of these has additional properties.

Wisniewski’s research on integration (Wisniewski & Middleton 2002) demonstrates that it is a very complex process, involving (at the least) alignment and construction. In alignment, “people identify correspondences between the modifier and head noun concepts that help indicate how the modifier property is to be integrated into the combination”. He gives the example of rose mushroom, typically interpreted as ‘a mushroom with thorns attached to its stem’ vis-à-vis porcupine mushroom, typically interpreted as ‘a mushroom with quills attached to its hood’. Clearly, the alignment in each ENC is different, and crucially dependent on the spatial configurations of modifier and head (and not on generic world knowledge). Moreover, alignment is insufficient, because a new version of the property must be constructed, the second stage of integration. Thus, the thorns and the quills on the mushroom are not equivalent to the thorns and quills on the rose and porcupine respectively. They have been assimilated into a new compound (a zebra horse does not have striped teeth or hooves, only striped skin). Wisniewski and Middleton (2002) confirmed this concept-specific alignment process by asking participants to draw pairs of ENCs such as helicopter flower and windmill flower and bucket bowl and coffee cup bowl. They consistently used specific spatial information about the nouns in their pictures (i.e., horizontal vs. vertical flowers, and handle on rim vs. handle on side, respectively).

3.2 The CARIN model

Christina Gagné and colleagues (Gagné & Shobin 1997; Gagné 2000, 2001, 2002; Gagné & Spalding 2004a, 2004b) have developed the Competition-Among-Relations-In-Nominals (CARIN) model of noun-noun comprehension. The CARIN model claims that “during conceptual combination, various relations become activated and compete with one another for selection”. In contrast to the Dual-Process model,
CARIN posits that property compounds are actually a specific type of relational compound, namely, a resemblance relation (e.g., a tiger snake is a snake that resembles a tiger) and so do not require a distinct process of interpretation. That is, a single relational process accounts for both types of nominal combinations. Gagné (2000) proposes a serial model in which the various relations (e.g., location, resemblance, etc.) are attempted in order of frequency; more frequent relations are attempted first, less frequent relations are attempted only later. Because resemblance is an infrequent relation, attributive combinations take longer to comprehend and are attempted after relational ones (Estes 2003).

As in the other accounts, the CARIN model seeks to account for the extra information in the system, again as an abstract relation between words in the mental lexicon. Two points are interesting about Gagné’s research in relation to this: first, the availability of particular relations with particular modifiers affects ENC comprehension (Gagné & Spalding 2004a). For example, mountain N activates a locative relation and chocolate N activates a made.of relation. Heads can also activate relations, such as N shop activates the sells relation. The ENC chocolate shop therefore triggers a competition between (at least) two relations ‘a shop made of chocolate’ and ‘a shop that sells chocolate’. Combinations where the frequently used relation does not apply, as in mountain magazine, are harder to interpret.

Secondly, Gagné has adduced evidence for relational priming. She demonstrates that student vote ‘vote by a student’ is processed faster after recent exposure to student accusation ‘accusation by a student’ than after student car ‘car for a student’. (Relational priming is also only activated for modifier relations; so chocolate shop fails to prime toy shop.) An area of continued interest to Gagné and others is the trigger(s) for relational priming. Gagné (2001) concludes that “research with novel compounds has shown no such relation priming in the absence of a common constituent (Gagné 2001)”. That is, priming does not work by triggering an abstract [N by a N] relationship, but only by a [student by a N] relationship. Estes (2003), however, disputes this finding by specifically testing for abstract relational priming. Interpretation of rugby shoes was facilitated when preceded by swimming flippers but not by road construction. Estes argues that Gagné failed to observe such priming because she did not use precise-enough relations between modifier and head in priming. Estes concludes that “the conceptual relation is an independent representational structure” (Gagné & Spalding [2004a, 2004b] seem to concede Estes’ view). The proposal in Section 4 offers a specific proposal to explain priming.

3.3 The C³ model

Fintan Costello and Mark Keane (Costello & Keane 1997, 2001) have developed the Constraint Theory of Concept Combination (C³) model of noun-noun comprehension. “Constraint theory has three components: (i) a generative mechanism to produce candidate interpretations; (ii) three constraints – diagnosticity, plausibility, and informativeness – that decide between these interpretations and (iii) a set of knowledge as-
sumptions about the types of concepts used in the combination”. The third component of C\textsuperscript{3} contains complex lexical knowledge.

What is interesting in the C\textsuperscript{3} model is its three constraints of well-formedness, which make it a strongly pragmatic theory:

i. The \textit{diagnosticity} constraint favours interpretations that contain strongly distinguishing properties of both concepts. Thus \textit{cactus fish} as ‘a prickly fish’ is favoured over ‘a green fish’ because \textit{prickly} is more diagnostic of \textit{cactus} than \textit{green}.

ii. The \textit{plausibility constraint} favours interpretations that describe the most plausible objects (with respect to physiology, geography, biology, etc.). So \textit{angel pig} is more plausibly a pig with wings on its torso than on its tail (\textit{ergo} our \textit{tree elephant} is gravitationally challenged!). Likewise, few people would expect \textit{tree crab} to have a standard \textit{animal (location)}\textsubscript{loc} semantics, because an aqueous habitat is expected. (Incidentally, an animal called \textit{the tiny Jamaican tree crab} really does spend its entire life in a tree!)

iii. The \textit{informativeness constraint} requires that compounds communicate something new. This explains the unacceptability of \textit{pig pork}, \textit{cow beef}, and so forth. Only one possible example of a non-informative (\textit{tiger/tree/elephant}) ENC, \textit{tiger stripes}, was found in the BNC corpora:

(5) Luce became aware of sunlight lying like a warm caress across her closed lids, and opened her eyes to find the shutters throwing \textit{tiger stripes of light and shade} across the bed. (BNC JY2 00839 = \textit{Joy bringer}. Wilkinson, Lee, Mills and Boon, Richmond, Surrey [1992]).

Costello and Keane (2001) illustrate the difference between C\textsuperscript{3} and the Dual-Process models with the example of \textit{bumblebee moth}.

C\textsuperscript{3} favours interpretations (a) and (c), because they have greater diagnostic features but Dual-Process favours (a) and (b), because they have more easily alignable features. Wisniewski (2001; cf. Wisniewski & Middleton 2002), however, points out that (c) \textit{is} alignable in that moths and bees have tapered torsos. Thus the sting on a bumblebee’s tail can align with the same spatial location on the moth, where no sting exists. (He also points out several misunderstandings of the Dual-Process model as presented in Costello & Keane 2001.) He concludes that alignment is always pertinent in interpretation of ENCs.
3.4 Summary

From the three theories surveyed above we can observe that psycholinguistic considerations of ENC comprehension significantly broaden the range of information to be accounted for: beyond any doubt, readers/hearers of ENCs draw on a wealth of information to formulate their interpretations, starkly evidenced by the perceptually-determined differences between *rose mushroom* and *porcupine mushroom*, and the alignable torso-feature of *bumblebee moth*. In addition, not all information is equally salient, as the CARIN model demonstrates. In addition, the ENC itself activates semantic possibilities (*chocolate X = X made of chocolate*) which can also prime subsequent ENCs (*student accusation* primes *student vote* more than *student car*). The linguistic question posed by Section 3 is, how can we account for the linguistic data (Section 2) in a way that coheres with the psycholinguistic results (Section 3)? To this question we now turn.

4. The ENC construction

The most important point is that constructions are nothing more or less than patterns of usage, which may therefore become relatively abstract if these patterns include many different kinds of linguistic symbols. But never are they empty rules devoid of semantic content or communicative function. In usage-based approaches, contentless rules, principles, parameters, constraints, features, and so forth are the formal devices of professional linguists; they simply do not exist in the minds of speakers of a natural language. (Tomasello 2003: 100)

Sections 2 and 3 reviewed the linguistic and psycholinguistic accounts of ENCs respectively and demonstrate that a whole raft of conceptual/semantic, lexical, syntactic, prosodic, and contextual information must be brought to bear on their successful communicative use. The question of how to account for the various data that belong together for ENCs of course concerns all grammatical constructions. It is the particular challenges of ENCs that make them a fertile field in which to explore plausible answers to this question.

So rather than going straight to ‘Now we will analyse construction X using theory Y’ we will linger and consider this broader question more carefully, namely, what should a grammatical theory look like if it is to account for this *belonging together* of related information? After all, belonging together characterizes human speech from its inception: “Functionally speaking, children’s early one-unit utterances are entire semantic-pragmatic packages – holophrastic expressions – that express a single relatively coherent, yet undifferentiated, communicative intention” (Tomasello 2003: 39).

The following section (4.1) is a condensed version of the constructional answer to this broader question; fuller versions are found in Goldberg (1995), Croft (2001), and particularly Tomasello (2003). Section 4.2 returns to the ‘analyse construction X using
theory Y’ manoeuvre in a way that further motivates the constructional approach while cohering with the data of Sections 2 and 3.

4.1 Construction Grammar

The basic idea of Construction Grammar (CG) is as follows (Croft & Cruse 2004; Croft 2005; Fried & Östman 2005): grammatical knowledge is organized in conventionalized self-contained routines of language usage in which all the information needed to execute the routine successfully – semantic, syntactic, phonological, pragmatic, conceptual, and so forth – forms a singular complex unit, namely, a construction. Prime examples are idioms, such as kick the bucket, which have information such as: morphosyntax: [kick–tense the bucket], semantics: die; interpersonal: used in non-formal communication, and so forth. Constructions are gradable in terms of schematicity: highly schematic constructions, such as generic English noun phrase (syntax: [DET N]NP; semantics: denote an entity; interpersonal: refer to an entity) are more abstract generalizations over less schematic, more contentful constructions, such as, for instance, a subclass of ENCs (syntax: [a [N HABITAT N ANIMAL]NP; semantics: an ANIMAL that lives in a HABITAT, etc.). Because constructions range from individual morphemes to clauses and beyond, all grammatical units are symbolic, specifying the form and function of the unit. Furthermore, constructions are in a network of relations: [a [N HABITAT N ANIMAL]NP is not only less schematic than [DET N]NP, they are both related constructions taxonomically (Tomasello 2003:315–320).

But what is the Sprachphilosophie of CG? How is it different from the rules-and-exceptions, algorithmic approach of other grammars? This is how we can characterize it: linguistics has a long tradition of approaching the problem of language in the way all large problems are dealt with, by dividing them into smaller problems. The scalpel of choice generally has been to incise horizontally across the pathway from conceptualization to expression, resulting in modularized strata for each stage: cognitive, pragmatic, semantic, syntactic, phonological, prosodic, and so forth (Seuren 2004 in response to Croft 2001). Once each module is described, algorithmic links are forged between modules, resulting in a production-line view of language (Levett et al. 1999). Such a divide and conquer strategy, in the view of many, comes unstuck when considering “peripheral” phenomena – idioms, expletive-infixation (un-bloody-believable), stranding errors (he catted the pat), and the like – which are all characterized by the need for a module to “look ahead” in order to use otherwise inaccessible information further down the production-line (Croft 2004 in response to Seuren 2004). In other words, a horizontal theory-based division of labour results in dispersal of information that needs then to be recollected in various ways for the purposes of usage.

5. Modularists solve this by (i) introducing feedback mechanisms, whereby a process waits while a look-ahead routine gathers information to feedback into the process; (ii) using the lexicon to store all potential look-ahead information; and/or (iii) allowing for parallel processing, having several production-lines running at once.
CG, which arose precisely from consideration of idioms, in contrast, makes vertical incisions through the system that produces natural clusters of (cognitive, semantic, syntactic, etc.), information, each cluster being a construction. Knowledge of a language is knowledge of how to successfully traverse successfully from conceptualization to expressions and back. The question of which pathways constitute a construction and not just an ad hoc utterance is answered largely by usage-based data – a description of a language is precisely a description of those conventionalized (entrenched) constructions that are shared by a speech community. In contrast to the horizontalists, a vertical usage-based division of labour results in dispersal of information that needs to be recollected in various ways for theoretical purposes. In CG the recollection results in a “network” (or “hierarchy”, “lattice”) of constructions, because otherwise the necessary horizontal generalizations are missed. For example, \([N N]_N\) is one pathway within \(\text{det} [N N]_N \text{NP}\), which is one pathway within \([\text{SUBJ}, \text{det} [N N]_N \text{NP} \text{ InTrans.Verb}] \text{CLAUSE}\), and so forth.

In other words, the two approaches use different criteria to guide the slicing: horizontal slicing is guided by similarity, encapsulation, and economy – for example, NPs, VPs, and DPs have similar properties and have a level of separateness (autonomy, independence) from syllables, nuclei, and metrical feet. The best theory has the fewest variables and most efficient algorithms; vertical slicing is guided by execution, integration, and sharedness – to execute successfully an utterance (i.e., to communicate) in a speech community all the necessary information must be learned, stored, and used together as a construction; it must be integrated. The best theory accounts for all the constructions that are shared by a speech community, an important indicator of sharedness being frequency of occurrence ( Tomasello 2003:169–175).6

Finally, it can be observed that although CG often presents constructions as form-function pairings the above summary suggests cognitive–expression pathways is perhaps a better characterization. Along the pathway exist various frames (i.e. forms) with slots to be filled by varieties of information, the whole construction serving a variety of functions.

Tomasello (2003; cf. Goldberg 2003) also argues persuasively that the constructional account makes the best (and in places, the only) sense of language acquisition, as the quote at the start of Section 4 shows. He amasses a wealth of empirical data to support this claim.

The following proposal, accordingly, is a constructional one in that the ENC pathway from conceptualization to expression and back is sketched. As a preliminary sketch, let us recall Table 5 above, as a suggested means of abstracting across three linguistic accounts of ENCs.

The question before us then is, if conceptualization of a certain entity in a certain way provides English speakers the option for expression of this as an ENC, what are the

6. The two approaches also posit different solutions to the dispersal/recollection problem.
frames along the pathway that are filled in the process of executing the ENC routine? To answer this I shall use a thought experiment: how would someone interpret tree flunq if told that a flunq is similar to a goanna, an Australian tree-dwelling reptile? The thought experiment rests on the assumption that two (or more) constructional pathways overlap at the semantic level but differ at the expression level. It is the rapid traversal and interactions of these two pathways that is of interest.

Section 4.2 considers the relational ENC tree flunq and 4.3 extends this to property ENCs, using tiger clam as a test case. Section 4.4 briefly looks at uninterpretable ENCs and after a summary in Section 4.5, Section 4.6 demonstrates the psychological adequacy of the proposal.

4.2 Relation ENCs

Suppose I explain to someone that a new species of reptile has been discovered called a flunq. It is similar to the goanna. In fact, two varieties of flunqs were discovered, one of which has been named a tree flunq. I then ask that someone at that point, What do you think a tree flunq is? The answer, as expected, would presumably be a flunq that lives in a tree. How is it that with no prior knowledge of flunqs, the most probable definition is given, one that is semantically equivalent to tree flunq and that most if not all people would agree with? That is, how do we explain the rapid mapping from two different syntactic structures to a single semantic structure? This mapping can be represented, with a schematic semantic representation as follows. (It should be noted that examples (6)–(7) and the following examples, represent the shared semantic layer of two overlapping constructions. The ENC construction is the combination of only the left-hand side of the syntax layer with its links to the semantic layer. This represents the assumption mentioned above.)

7. Note, the focus is only on the semantic and syntactic components of the construction. A further improvement would be to specify prosodic information to disambiguate the relational and property readings of an ENC, such as baby photographer, ‘a baby-like photographer’ versus baby photographer, ‘a photographer that photographs babies’.

8. I adopt the following conventions: specific lemma: \textit{lemma}; specific lemma/operator as a predicate: \textit{pred}; semantic primitive: \textit{animal}; lemma member of semantic primitive as a predicate: \textit{animal’}. 

---

Table 5. Constructional construal of Levi, Fabre and Sébillot, and Ryder

<table>
<thead>
<tr>
<th>Theorist</th>
<th>Empty</th>
<th>Saturated</th>
<th>Expressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levi</td>
<td>N cause N</td>
<td>stress cause heat stress</td>
<td>‘thermal stress’</td>
</tr>
<tr>
<td>Fabre and Sébillot</td>
<td>lemma \textit{(lemma\textit{N-Rule})}</td>
<td>cover (earthquake\textit{theme})</td>
<td>‘earthquake coverage’</td>
</tr>
<tr>
<td>Ryder</td>
<td>x hunt y</td>
<td>tiger hunt man</td>
<td>‘tiger man’</td>
</tr>
</tbody>
</table>
One explanation is that there is a generic equivalent of (6), something like (7):

\[
(7) \quad [a \text{ } N_{\text{HABITAT}}, N_{\text{ANIMAL}}]_{\text{NP}} = [a \text{ } N_{\text{ANIMAL}} \text{ } [\text{that lives in a } N_{\text{HABITAT}}]_{\text{PP}}]_{\text{REL}}_{\text{NP}}
\]

What (7) indicates is that the two speakers share the same conceptualization, for our purposes represented as habitual’ (live’ (ANIMAL’)) (HABITAT’)_loc, which the first speaker expresses as a tree flunq and the second speaker, prompted by the elaboration request, What is a tree flunq?, answers felicitously as A tree flunq is a flunq that lives in a tree. We have in this transaction bidirectional traversal of the ENC construction, from conceptualization for speaker A and to conceptualization for speaker B.

The justification for (7) is as follows: imagine the situation if there was no such pre-existing, generic construction as (7). Any alternative must explain where the habitual’ (live’ x, y) and the loc come from, because it is the predicate and thematic roles between the arguments that is crucial to the composition. The most viable alternative to (7) is to posit on-the-fly construction based on alignment of lexical information concerning both terms. But this is problematic for three reasons.

(i) The information required for both lexical entries would be considerable. For tree we need to know that although it is not a place, it does contain “places” where animals can live habitually. We need to know information about the weight-bearing capacity of such places. We need also to know what a prototypical tree is so that we do not include inhabitable trees like bonsai trees. For flunq we need to know its approximate weight, that it is an animal and has sufficient biological features that allow tree-dwelling, such as mobility, balance, weight, and so forth. (That is, the C³ plausibility constraint is activated.)

However, all this information is insufficient of itself to select the appropriate HABITAT usage of tree and construct the appropriate relationship with flunq. This combining occurs through considerable inferential processing that decides that flunqs in all probability should be able to live in trees. How then, does the rapid translation into ‘a flunq that lives in a tree’ occur? Would it not be more efficient to have a mechanism for linking generic animals and generic habitats which alleviates the need to carry out complex inferring every time?

(ii) The word flunq is only recently stored in the mind of the hearer and yet he/she responds immediately (at least, that is a testable hypothesis as Wisniewski [pc] pointed out). How is all this lexical information concerning flunqs derived? The lexical-only explanation must copy in toto the entry from goanna to flunq and then perform the in-
ferences on the new information. Again, this hardly seems cognitively or theoretically parsimonious.

(iii) If the linkage is based on cognitive processing from two encyclopedic-type entries in the mental lexicon, no interpretation should arise for tree elephant, as no links can be made. But ask any eight-year old what a tree elephant is, and they will reply ‘an elephant that lives in a tree’ (on the assumptions that [i] they know the ENC construction and [ii] their plausibility constraint is less encumbered by adult predilection for realism!). So the problem with tree elephant then is its physiological implausibility, which emerges after the semantic composition of the following:

(8) habitual´ (live´ (elephant´) (tree´),loc)

A lot hinges on this ‘after’—it entails that a generic construction such as (7) must have been utilized. One could argue that the compound itself, tree flunq, establishes a generic relationship (akin to Downing’s [1977] generic related.to’ predicate), but this is still an argument for a construction such as (9) below. The only difference is the semantic granularity of (7) vis-à-vis (9).

(9) [tree flunq]N = [flunq |that V |PP a tree|REL]N

But (9) would not solve the tree elephant problem. Why is ‘an elephant that lives in a tree’ the rendering of this if the construction is generic in the way (9) is? Again, the most plausible explanation is that there is a generalized construction containing semantic primitives to which specific (i.e. activated) lemmas are associated.9

4.3 Property ENCs

Can this reasoning be applied to property compounds, such as tiger clam? For the reasons given above, mutatis mutandis, I think it can. However, a lot depends on getting the basic construction right for the addressee.10 What tiger does is to assign some feature of the modifier to the head, giving it a feature that by implication (>) it does not already have (that is, the criterion of informativeness; for example, tiger zebra is less informative than dwarf zebra because zebras have stripes, a strongly alignable and diagnostic property of tigers). The construction of tiger claim is something like the fol-
lowing, where $Q_{\text{lemma}}(\text{pred}')$ means pred’ is a property of lemma (the $Q$ is meant to suggest Pustejovsky’s [1995] qualia structures) and $\neg = \text{not}:

\begin{equation}
(10) \quad [\text{tiger clam}]_{N} = [\text{striped clam}]_{N}
\end{equation}

\begin{equation}
\text{have'}(\text{clam'})(Q_{\text{tiger}}(\text{striped'})) > \neg Q_{\text{clam}}(\text{striped'})
\end{equation}

\[\text{syntax}\]
\[\text{semantics}\]

What generic construction, therefore, can we posit for such animal animal property compounds? Note that the striped’ predicate is lemma-specific (proof: tiger clam as ‘a ferocious clam’), so it cannot be replaced with a semantic primitive. I suggest, therefore, the construction in (11) for animal animal compounds.

\begin{equation}
(11) \quad [N_{\text{ANIMAL}}^1, N_{\text{ANIMAL}}^2]_{N} = [A_{\text{PRED}}, N_{\text{ANIMAL}}^1]_{N}
\end{equation}

\begin{equation}
\text{have'}(\text{ANIMAL}'')(Q_{\text{ANIMAL}}(\text{pred'})) > \neg Q_{\text{ANIMAL}}(\text{pred'})
\end{equation}

\[\text{syntax}\]
\[\text{semantics}\]

The justification for including the implication in the construction is that knowledge of the construction includes knowledge of the implication. (Such a claim is a testable.) Moreover, this representation alerts us to a potential problem in coercing property compounds into relational compounds by positing a \([\text{resemble}'(\text{ENTITY}'')(\text{ENTITY}'')_{\text{THEME}}] \Leftrightarrow [N^2 N^1]_{N}\) construction, as proposed by Gagné (2000; and disputed by Estes 2004:308, 315), because relational compounds do not have a $> \neg Q_{\text{ENTITY}}(\text{pred'})$ implication.

Therefore, the proposal can be extended to include property compounds.

4.4 Opaque phrasal ENCs

Finally, it is interesting to consider phrasal ENCs that fail to make sense according to the speaker/listener. Because the construction is headed by the second noun,\(^{11}\) it nevertheless does fit into a superordinate \([\text{MODIFIER NOUN}]_{N}\) construction,\(^{12}\) in which case the modifier can be treated as an uninterpretable label.

To illustrate, consider elephant garlic (Allium ampeloprasum), a large variety of garlic/leek, with a somewhat sweeter taste than regular garlic. Let us assume that the hearer cannot construe ‘a large garlic’ or ‘a garlic that elephants eat’ or any other read-

\[\text{11. Moder (2004) discusses examples such as traffic nightmares (accent MPA), which seems to refer to traffic and not nightmares (compare traffic nightmares). But nightmares is a metaphor here for problems, and it is to the problems (traffic ones) that the ENC refers. So we can maintain the headedness of the second noun in all cases.}\]

\[\text{12. We could also say it overlaps with this third, superordinate generic pathway, whose level of activation may increase in reaction to failed traversals of other ENC constructions. Thanks to William Croft (pc) for clarifications on this section.}\]
Tree tigers and tree elephants

ing. It would seem highly unlikely, however, that they could not construe, ‘an elephant garlic is a type of garlic’.

What is significant (structurally and cognitively) in this rendering, is that the construction effectively lexicalizes the compound to force a viable interpretation, by treating the ENC as a non-decomposable entity (i.e. ENTITY\(^1\)-ENTITY\(^2\)), blocking relational and property construals. Thus I propose an additional construction, the type-denoting ENC construction, whose cognitive activation might be a last-resort strategy, or might in fact be ubiquitous but backgrounded.

\[
\text{syntax} \quad \text{semantics}
\]

\[
(12) \quad [N^1 N^2]_N = [[N^1 N^2]_N [\text{that is a type of } N^2]_{REL}]_N
\]

4.5 Overview

Granting that the above thought experiment is accepted, the constructional account of ENCs can be presented, using the production and comprehension of tiger clam with a property interpretation.

First, diagrammatically, tiger clam can be represented as in Figure 1 (omitting \(\neg Q_{\text{lemma}}(\text{pred}')\)). The lexicon is separate in both cases because it is like a rucksack for the journey to and from conceptualization to expression, the place to store and retrieve information as need be, a grammar in miniature. The asterisk represents an empty frame.

Secondly, lexical entries contain two types of information: information specific to that entry only (e.g. unicorn has \textit{tapered horn on snout} and \textit{imaginary animal} in its entry) and links to semantic primitives (e.g. unicorn is an \textit{animal}'(\textsuperscript{e}). This provides a natural set of diagnostic properties for alignment, because only unique information, such as \(Q_{\text{tiger}}(\text{striped}')\), is listed.\(^{13}\) The use of primitives in lexical entries eliminates unnecessary redundancy of information and allows lexical entries to be constrained to only contain information specific to that entry. It is this information that is used in processing of ENCs to search for the best interpretation. This processing specifies the final interpretation (i.e. Wisniewski’s alignment process) in ways that the generic constructions cannot.

Third, the semantic primitives in turn appear in constructions, acting as filters for lexical content. The primitives therefore provide the integrating link between lex-

\(^{13}\) Non-diagnostic properties, such as \textit{four-legged}' are located in the generic entry of (a sub-domain of) \textit{animal}', namely \(Q_{\text{animal.mammal}}(\text{four-legged}')\). In other words, primitives themselves are lexical entries.
icon and semantics/syntax. Constructions are organized taxonomically in much the same way as the lexicon, in a network that Goldberg (2003) calls a ‘construct-i-con’. At the top are the most generic constructions, in our case [N N]N, branching (perhaps with intermediate groupings) into a finite number of [N entity N entity]N constructions, branching further down (from any higher node) if necessary into more specific hyponymic constructions, a stellar example being the [N entity N]N construction that has a specific lemma in place.

Fourth, from a production point of view, once the semantics is composed, if a lemma is activated, such as tiger, its semantic primitives are activated (e.g. ANIMAL) which in turn activates the appropriate construction(s) containing ANIMAL.

Finally, from a comprehension point of view, the construction tiger clam, activates the lemmas tiger and clam so that their parts of speech (nouns) and primitives (ANIMAL) are activated. This allows syntactic parsing of [N entity N entity]N, which triggers all constructions containing this configuration, including have’ (ANIMAL) (QANIMAL (pred’)), which gives rise to the interpretation have’ (clam’) (Qtiger (striped’)). Cognitive processing then works out the details: How are properties related/aligned? Does the interpretation satisfy the three constraints of diagnosticity, plausibility, informativeness? Does the interpreter need to look further back in the discourse context

---

14. In distinction from projectionist models such as Role and Reference Grammar (Van Valin & LaPolla 1997), which map directly from lexicon to semantics without any such primitives.
for processing clues? Compounds that fail to retrieve primitives, such as onij bird, or primitives with no corresponding construction, such as sound silence, trigger the sortal ENC construction (which contains no primitives).

4.6 Psychological adequacy

This section demonstrates that the proposed ENC constructions account for the range of phenomena outlined in Section 2.

4.6.1 The Dual-Process model

The construction clearly coheres with a weaker version of the Dual-Process model, in which some ENCs trigger two interpretative strategies. This occurs when an ENC activates different constructions with identical syntactic sides (i.e., the reverse of (9), (10), and (11), where two constructions share the same semantic side). Consider, for example, hippopotamus worm, which could be property-assigning, 'a large worm' or relation-assigning, 'a worm that lives in a hippopotamus.' The constructional analogue of the Dual-Process model is as follows, where (inside a) hippopotamus is taken as a habitat:

\[(13)\]

The model is compatible also with a stronger version of the Dual-Process model in which all ENCs trigger at least two interpretations. For instance, tiger clam, of which a relational interpretation is strained to say the least, might also trigger a creative (!) relational interpretation through metaphor (an animal is a habitat), resulting in \[N_{\text{animal}} \rightarrow \text{habitat} N_{\text{animal}}\text{'s} N\], 'a clam that lives in a tiger'.

4.6.2 The CARIN model

The competition among diverse interpretations of relation ENCs as posited by CARIN occurs whenever an ENC fits into two construction templates. For example, Warren (1978) discusses the two readings of brick warehouse, 'a warehouse that is made of bricks' and 'a warehouse that contains bricks'. As noted in Section 2.3, despite being

---

15. Gagné and Wisniewski (and Estes) seem to have different interpretations of the Dual-Process model. Gagné (pc, her emphasis) writes, "the theory proposes that either relation-linking is used or property-mapping is used". But Wisniewski (pc) suggests that both processes occur only if both relational and property interpretations are possible (and, I would infer, necessary, in the case where one does not succeed before the other finishes). But as argued in the discussion of tree elephant in Section 4.2, plausibility adjudicates the viability of an interpretation, implying that both processes must always be activated. Perhaps, then, the procedural details of this model could be clarified.
categorized as semantic, Warren’s approach is essentially lexical. This is clear in her explanation of the two readings of brick warehouse. Different lexical features of brick and warehouse are “highlighted” in each reading. ‘A warehouse that is made of bricks’ highlights brick = material and warehouse = man-made object, and ‘a warehouse that contains bricks’ highlights brick = man-made object and warehouse = place. In the approach proposed here, we could recast this as two construction types:

(14) a. \[N_{\text{material}} N_{\text{container}}\]N ⇔ made.of´ (container´) (material´)
   b. \[N_{\text{artifact}} N_{\text{container}}\]N ⇔ contain´ (container´) (artifact´)

Both constructions are activated because brick links to material and artifact.

The proposal explains the relational priming effect (Gagné 2001; Gagné & Spalding 2004b; Estes 2003). The relation is primed by the semantic side of the construction, as for instance in example (7): habitual´ (live´ (animal´) (habitat´)loc) ⇔ [N N]N.16 Estes (2003:315) argues that “the conceptual relation is an independent representational structure”. The generic semantic frames of ENC constructions in the proposal are just that, independent structures indicating the conceptual relation.

Generic relational priming is not incompatible with lemma-specific relational priming; it would be largely a matter of conventionalization according to CG. In general, particular lemmas will trigger particular constructions on a regular basis. For example, chocolate has in its entry a link to the primitive food. Food in the construction (15) below; hence, chocolate triggers made.of relations.

(15) \[N_{\text{food}} N_{\text{entity}}\]NP = \[N_{\text{entity}} [that \ is \ made \ of \ N_{\text{food}}]\rel\]NP

Specifically, lemma-specific priming might be evidence of entrenched lemma-specific constructions, such as [det [chocolate Nentity]\]NP ↔ [a Nentity [that is made of chocolate]\rel]\NP.

The proposal therefore coheres well with Gagné’s CARIN model.17

16. The semantic granularity of the primitives in these constructions is a matter of ongoing research, as evidenced in Estes (2003) critique of Gagné in which it is alleged that she failed to control for the precise relationship in priming experiments. Section 4.7 on prototypes has suggestions on the granularity issue.

17. Gagné (pc), however, raises some questions about my proposal, concerning alignment in the Dual-Process model: “…how can one decide which alignable feature is relevant? …how does the system decide which property to use!” and concerning selection in general: “…when there are multiple constructions, what determines which of these is selected?”

To the second question, I would say that Gagné’s own work on discourse context provides the basic answer (Gagné & Spalding 2004a). They write (p. 454):

The overall picture of recent research on conceptual combination, then, suggests that the critical issue is the availability of interpretations. These interpretations might be
4.7 For further investigation

The preceding constructional account of ENCs makes extensive use of generic frames, arguing that they provide a central role in facilitating rapid production and comprehension. Extending the pathway metaphor (Section 4.1), we can say that between the highways ([det [N N]NP]) and goat tracks (a Claysons NENTITY) of ENC constructions, lie the avenues of generic constructions (det [NAnimal NAnimal]NP). These generic constructions carve out families of constructions, related on the basis of common membership of primitives associated with specific lemmas.

Such generic frames have empirical evidence – recall from the previous section (4.6.2) that Estes (2003:315) argues that “the conceptual relation is an independent representational structure”. They also have the support of William Croft (pc), who

differentially available due to relation frequency, to recent exposure to combinations sharing the modifier and relation, and to discourses that contain referents that embody the interpretation.

My proposal only adds a fourth factor: the context might provide activation of primitives that the terms in ENCs are connected to. For instance, in the sentence “One of the world’s most dangerous creatures is the tiger shark”, dangerous activates ferocity (and not striped!) and so the mention of tiger, linked to both ferocity and striped, activates the ferocity property. Modérs (2004) article on context provides further illustrations of contexts that do not necessarily contain “referents that embody the interpretation” but clearly predispose the hearer to posit particular interpretations.

To the first question, the best answer I can think of concedes a weakness in the proposal – a weakness of most linguistic theories – namely, its reliance on amodal semantic structures. Barsalou et al. (2003) review research into conceptual systems and conclude that modality-specific (i.e. perceptual, sensorimotor) information is essential in any conceptual knowledge system. They provide the following evidence: (i) perceptual information affects taxonomic and/or thematic activation, as when participants produce roots less often for lawn than for rolled-up lawn; (ii) modality-specific priming occurs, as when loud is verified faster for blender (auditory) when preceded by rustling for leaves (auditory) than by tart for cranberries (gustatory); (iii) texts implying perceptual orientation lead readers to simulate objects in that orientation, so birds in flight leads to faster processing of birds with outstretched wings than with folded wings; (iv) embodiment effects are well-documented, such as processing baby is faster when performing arm approach movements (pulling) than avoidance movements (pushing); and (v) domain-specific deficits occur in lesion patients. Modal knowledge in fact is evident in the review of the Dual-Process model, as clearly visual information is affecting noun-noun comprehension in cases such as porcupine rose vs. mushroom rose.

Accordingly, tiger clam might be best explained as an (imagined) visual blend between an image of a tiger and an image of a clam, the property selection being guided by the visually salient features. In this case there might be no need to posit amodal blends like have’ (clam’) (Qtiger(striped’)).

Estes (2003) conducted two experiments, summarized as follows: in Experiment 1, target combinations were more likely to be comprehended, and were comprehended more quickly, when preceded by prime combinations that used the same attribution or relation, despite the
in reviewing an earlier draft of the present article, wrote, “I don’t doubt that there are certain very widespread yet relatively specific semantic schemas that speakers draw on, such as Material.Noun + Object.Noun, to denote an object made of that material. But it’s probably going to be a matter of frequency...than a finite set of semantic primitives”.

One crucial question is to determine which generic constructions are entrenched on the clines of constructional schematicity? For example, for tree flunq, we can posit the following cline:


The italicized [HABITAT ANIMAL] generic construction was used in Section 4 and does a lot of work in the theory. Further empirical investigation and linguistic analysis is required to find evidence for and details about such generic constructions.19

This qualification notwithstanding, the proposed constructional approach to ENCs presented in Section 4 has a high degree of psychological adequacy and coheres well with Estes’ (2003:317) summary:

... research on nominal combination is beginning to converge on a general model. It may well be that attribution often entails comparison and mapping (Wisniewski 1997), which is driven by salient or diagnostic properties of the modifier (...Costello & Keane 2001...), and that relational comprehension entails thematic role-filling (Wisniewski 1997), which is driven by the prior use of particular conceptual relations (Gagné & Shoben 1997). Finally interpretations produced by these process are undoubtedly subject to the satisfaction of pragmatic constraints (Costello & Keane 2000...).

absence of lexical overlap between prime and target combinations (e.g. onion tears was primed by ocean nausea). Comprehension of mountain snake was facilitated by the prior presentation of jungle bird, despite the absence of lexical overlap between prime and target combinations. Experiment 2 tested to see if a particular relation or attribute was being primed or just relational or attributive processes in general. Experiment 2 held constant the process engaged by the prime and target combinations (i.e. attributive [=property] process or relational process), and manipulated whether a particular attribution or the particular relation was repeated from prime to target. For instance, the target copper horse was again comprehended faster and more easily when preceded by glass rose, which uses the same relation, than by sign post, which uses a different relation, even though prime and target combinations engaged the same process (i.e. relational combination). Thus, Experiment 2 demonstrated that the prime combinations activated particular attributes and relations, rather than the attributive process or the relational process more generally.

19. And for the relative ordering on the cline between pairs such as [place animal] < [habitat entity]; [habitat lizard] < [tree animal]; and [tree lizard] < [habitat flunq].
5. Conclusion

... putting forward a theory is like taking out a loan, which must be repaid by gleaning an empirical basis for it; theories that fail to do so... are declared bankrupt. In the sciences of the mind, this maxim translates into the need to demonstrate the psychological (behavioural), and, eventually, the neurobiological, reality of the theoretical constructs. (Edelman & Christiansen 2003)

The present article began with an overview of ENCs (Section 2) and then examined various psycholinguistic accounts of their properties (Section 3). In light of this, a constructional account of ENCs was developed (Section 4), modelled primarily on Construction Grammar, but supplemented with the use of generic constructions. These generic constructions, evidence for which has been provided by Estes (2003), are hypothesized to be entrenched in the cline of construction schemacity in a way that allows them to facilitate the rapid production and comprehension of ENCs in communication. The lexical primitives present in these basic constructions provide crucial links in the grammar, acting as customized filters through which cognitive/semantic content is passed. These generic constructions reduce complexity, act as turnstiles to related (overlapping) constructions, and rapidly increase processing. Their precise position on any one cline is a matter for further consideration.

Thus the maximum degree of psychological adequacy has been pursued, 20 a goal advocated by Butler (2003) as essential to the development of robust functional-cognitive theories of language. As Edelman and Christiansen (2003) state above, hopefully empirical testing would not declare this proposal bankrupt, even if its returns are lower than anticipated.

20. Janssen’s approach (this volume) could be understood as a critique of my CG approach (and hence its psychological adequacy), as he states, “Thus, no mental picture, conceptual, or semantic representation is assumed to be conveyed or portrayed by means of an utterance.”

But I do not see a fundamental incompatibility with our approaches: CG is amenable to an agentive interpretation à la Janssen, and Janssen’s approach does not preclude entrenchment of frequently-occurring linguistic acts with semantically stable “results of linguistic acts” (Janssen, this volume), results taxonomized by a symbolic representation.

My point in this article is that complex constructions that require high-processing load, such as ENCs (but unlike idioms), do not entail either radical semantic underspecification and/or a purely cue-based instructional-semantics (which Janssen advocates), nor do they entail the opposite approach of exhaustive lists of possible semantic relationships to be assessed algorithmically (which Janssen critiques). Rather, the shared communicative data, such as phonetic, contextual, gestural, visual, etc, is activated by and intends to activate for the interlocutors the appropriate constructions, constructions entrenched by frequent usage and replete with as much interpersonal, semantic, conceptual and other information as is possible to entrench, because ease and speed of processing – both pressing communicative goals – are facilitated by informational fixity. Thus CG has everything from radically underspecified to semantically detailed constructions (cf. the discussion of (a Clayton’s Nentity NNP in Section 4.5 above).
References


Tree tigers and tree elephants


English constructions from a Dutch perspective
Where are the differences?

Arie Verhagen
Leiden University

In construction grammar, specific constructions are daughters of more general patterns, the former inheriting properties of the latter, besides providing specifications of their own. Therefore, differences and similarities between languages may differentially involve lower and higher levels of generality. This paper demonstrates this by comparing three construction types in Dutch and English: the way construction, the time-away construction, and causative constructions. The first instantiates a productive pattern in English, but not in Dutch. The second is syntactic in English, and morphological in Dutch. The third inherits more properties from general clause syntax in Dutch than in English. The grammatical properties of each case nevertheless show important similarities. Low-level constructions may strongly determine cross-linguistic similarity, due to similarity of form-meaning pairing.

1. Introduction

In typology, terms like ‘analytic’ and ‘synthetic’ are quite commonly used to characterize a whole language, i.e. they are applied at the level of a complete grammar. One language may be said to ‘use syntax’ for the selfsame purposes that another ‘uses morphology’ for. In such statements, differences between languages are located at highly general dimensions of linguistic organization. Similarly, with respect to word order, a question like “SVO or SOV?” is often also asked at the level of the language as a whole, rather than at the level of a particular subset of expressions in the language. These tendencies reflect a view of languages as complete and coherent systems in which everything is connected to everything else; as is well known, this view goes back to (at least) Saussure’s view of a system of values that mutually determine each other, and thereby define the system as a whole.

A considerable amount of linguistic research is still directed at such very general properties as seem to be characteristic of complete grammars. But over the last two decades or so, parallel research into details of grammatical structure, actual language use, language acquisition, and language change has accumulated evidence for
the conclusion that all kinds of ‘in-between’ patterns of regularities exist; they have
to be assumed in order to account for all of language structure and processes of use
and change. A number of approaches have developed which incorporate this insight;
they differ in detail and sometimes also on more general points (cf. Croft & Cruse
2004:Ch. 10 and the references cited there; Langacker 2005), but they share the in-
sight that a substantial part of linguistic knowledge consists in knowing grammati
cal pairings of form and meaning – not only words and idioms, but also prototypical
‘constructions’: partly specified templates that have a conventional meaning, and one
or more open slots for variable constituents (The X-er, the Y-er; X[negative clause], let
alone Y[phrase]; What’s X doing Y?; let X Y[verb]; be Y[verb]-ing; etc.). The cover term
for these approaches is ‘construction grammar’.

In such a view, general grammatical rules (of the type [NP VP]) on the one hand
and specific lexical items on the other are considered to be limiting cases of ‘construc-
tions’ in an extended sense, with the prototypical constructions in the middle part
These constructions are related to each other in taxonomic networks, in the same way
as the mental lexicon is regarded as being organized in terms of conceptual taxonomic
relationships (hyponymy, besides possibly other kinds of connections), with the more
specific templates inheriting characteristics from the more general ones, without be-
ing completely reducible to such general patterns. In this way, we are witnessing the
emergence of a view of grammar as a much more loosely organized network of words
and constructions, at all kinds of different levels of abstractness, or schematicity. The
scientific virtue of generalization is no longer sought in general properties of the struc-
tural components of the systems as such, but in principles of organization (such as the
analogical and metonymic extension of prototypes), and especially in the way such
networks emerge from language use over time – both developmental time in indi-
viduals, and historical time in communities (Barlow & Kemmer 2000; Bybee 1985,
1995, 2001; Bybee & Hopper 2001; Croft & Cruse 2004:Ch. 11; Goldberg 2003, 2006;
Langacker 1987:Ch. 10; Tomasello 2003).

This view of grammar makes the kind of general statement that I mentioned at
the beginning rather suspect. Only if pieces of linguistic knowledge belong to neatly
distinguishable components of a grammar such as syntax or morphology is there any
reason to expect differences between grammars to observe boundaries between such
alleged components. In a constructional view, differences may be expected to occur
at any level of specificity, but not completely randomly: the hypothesis that the most
general rules of grammar are limiting cases of constructional templates predicts a cline
differences at different levels of generality, with a larger number of differences at
lower levels yielding greater differences in general rules. Moreover, differences may
also occur in the organization of the networks of constructions, and do not have to be
manifested (only) in the constructions themselves.

All in all, the point is that a constructional view of grammatical organization rec-
ognizes the important role of (historical) convention, since many properties of lower-
level schemas – constituting crucial components of a language user’s knowledge – are
not straightforwardly predictable from more general properties of the language. In this respect, this approach differs fundamentally from structuralist (including generative) approaches, which seek to ‘explain’ as many specific phenomena in a language as possible in terms of properties of the system itself.

Comparative research is especially useful here. Often, some sort of correlation between a more specific and a more general phenomenon can be established, and then there is a scientifically understandable temptation to explain one in terms of the other. But such a correlation within a single language is in itself only a very weak indication of its proper explanation. If one other language exhibits the same general property but lacks the more specific phenomenon, or the other way around, this suffices to show that a strict system-based explanation cannot be the whole story, and that at least some element of convention is also involved. What I will do in this paper is compare some aspects of English and Dutch that illustrate precisely this point. The fact that this can already be done relatively easily with such closely related languages provides additional support for the general idea that grammatical systems are not very tightly integrated, but organized rather loosely, with considerable space for variation in conventional solutions to similar communicative and conceptual problems.

First, I will discuss the similarities and differences between the way construction in Dutch and in English. At a relatively low level of analysis, they appear rather similar, but the English case seems much more integrated into the rest of the grammar of English – in the sense of sharing properties with other words and constructions, especially general and productive ones – than the Dutch one. Nevertheless, it is not at all clear that this seemingly better integration makes any difference for the status of the English construction in speakers’ knowledge of their language. Next, I compare the syntactic time-away construction of English with a particular morphological construction of Dutch, involving the prefix ver-, that has the same function as the syntactic construction has in English. The general difference between morphology and syntax turns out not to be very illuminating for understanding specific grammatical properties of the two constructions in the two languages. The English case even turns out to be more like the Dutch one than existing analyses of the English construction have suggested. Finally, I compare the possibilities for marking participant roles in causative constructions across the two languages. Here, it is Dutch that appears to exhibit a more tightly integrated system (in the sense indicated above), but here too, not much seems to follow from this difference for understanding properties of the specific patterns.

All data discussed stem from the studies cited. The empirical basis of this research involves corpora and informants, and in some cases grammars; for details, both of facts and of analysis, the reader is referred to the original publications.

2. The way construction

The way construction provides a typical and much discussed example of a grammatical template in English (Jackendoff 1990; Goldberg 1996). In Verhagen (2003), I presented
a (relatively detailed) analysis of its parallel in Dutch. English examples are given in (1)–(2) and Dutch ones in (3)–(4).

(1) Pat pushed her way out of the room.
(2) Volcanic material blasted its way to the surface.
(3) Zo blufte zij zich een weg uit Auschwitz.
   Thus bluffed she refl a way out Auschwitz
   ‘That’s how she bluffed her way out of Auschwitz.’
(4) Twee bussen boren zich een weg naar het hart van Istanbul
   Two buses bore refl a way to the heart of Istanbul
   ‘Two buses are boring their way to the heart of Istanbul.’

Such sentences exemplify a particular construction because they exhibit a number of systematic, correlated properties – both in form and in interpretation – which cannot be explained on the basis of their general grammatical structure (in combination with the ordinary meanings of the words): the subject referent creates a (possibly metaphorical) path and/or removes obstacles on it, and the subject referent moves along this path, also when the verb, e.g. push in (1), does not normally indicate either movement of the subject or the creation of something. Moreover, the nouns (way, weg) are necessary elements of the respective constructions. Thus, English has a template (informally) “to verb one’s way + oblique (locational) phrase”, and Dutch a template “zich een weg verb + oblique (locational) phrase”, each conventionally associated with a specific meaning.

The Dutch and English constructions explicitly differ in a number of respects (in particular, the use of a possessive vs. a reflexive pronoun), but two important differences are not visible in the templates themselves. The first concerns the fact that both English and Dutch have a default verb for this construction, but that this is not the same verb: English uses to make, Dutch banen. Thus, English has the specific pattern “to make one’s way through X” and Dutch “zich een weg banen door X” as specific templates, separately stored in speakers’ long term memories, but inheriting the properties of the more general ones. The functional similarity is that these verbs do not add meaning beyond the meaning of the construction. The difference is that make is a quite general verb, while banen is highly specific: in fact, the latter only occurs in this construction. Thus, while the specific pattern in English also inherits the properties of the well entrenched and productive verb make, its parallel in Dutch only inherits characteristics from the more general template; the difference is represented in the partial networks in Figures 1 and 2, respectively.

1. Plus, actually, one more that is closely related and resembles the historical source of the construction discussed in the text. In this related case, banen also obligatorily combines with the noun weg and has a largely overlapping meaning: it indicates the creation of a path and the removal of obstacles, but not necessarily movement along the path (although it may, depending on the verb). For details, see Verhagen (2003); for the diachronic development, Verhagen (2002).
Thus, although the English and the Dutch way constructions appear to have much in common in terms of structure and function, the way they are integrated into the overall networks of linguistic units is not the same in the two languages. The same in fact holds for their connections to more general constructional patterns (the parts indicated by “...” at the top of Figures 1 and 2). The specific pattern “zich een weg banen door X” (with both the verb and the path-marker lexically specified) as well as the somewhat more general superordinate schema “zich een weg +V + OBL” have to be conceived of as stored in the long-term memory of speakers of Dutch, with the specific schema as the prototype of the general one. Analogously, the prototypical pattern “to make one’s way through X” and its superordinate “to +V one’s way + OBL” are stored in long-term memory of speakers of English.

But even at a level that is only slightly more abstract, differences emerge. The English way construction has been characterized as a specific case of resultative constructions (of the type He cried his eyes red, so-called fake-object resultatives; cf. Goldberg 1996:50, and references cited there, for discussion). It exhibits a transitive pattern, with two argument positions (subject and object). But the Dutch way construction, with its characteristic reflexive element zich, exhibits a ditransitive pattern, with three arguments: subject, direct object and indirect object; it actually looks like a kind of benefactive construction. So for Dutch, the extended taxonomic network should, as it turns out, look like Figure 3.
In English, however, the more general pattern to which the minor network of \textit{way} constructions should be subordinated is the transitive one, as it is a kind of resultative; i.e. it should look as in Figure 4.

What this suggests is that the position of these constructions in the ‘grammatical space’ of Dutch and English is quite different for each language. However, it should be noticed that the benefactive pattern near the top in Figure 3 is not at all a productive pattern in Dutch. There is a conventional, productive pattern “\textit{zich een weg +V}”, which can be glossed as “to \textit{V oneself a way}” and which (roughly) means “create a path/opportunity for oneself (and use it), by means of \textit{V-ing}”, but it is \textit{not} an instantiation of a more general pattern “\textit{iemand +Y +V}” (“to \textit{V someone Y}”), meaning “to make Y for someone by \textit{V-ing}”. Curiously enough, English does have a productive pattern of the latter sort; while (5) is unacceptable in Dutch (this can only be expressed as in (7)), the English parallel (6) is perfectly acceptable.

(5) \textit{Jan maakte haar een boterham.} 
(6) John made her a sandwich. 
(7) Jan maakte een boterham voor haar. 
\hspace{1cm} John made a sandwich for her 
\hspace{1cm} ‘John made her a sandwich.’

There is something of a paradox here. English has a fairly productive general benefactive construction, but the \textit{way} construction is not an instance of it – Dutch does \textit{not}...
Figure 5. The Dutch \textit{weg} construction as an island

have a productive benefactive construction, although its way construction does seem to instantiate it. In any case, the consequence is that we have to exclude the way construction from the network of Dutch ditransitive constructions, and replace Figure 3 by Figure 5: in Dutch, the network of more and less specific way constructions actually constitutes a kind of island in the whole of the grammar.

In English, on the other hand, the way construction does not constitute such an island. Although it has a number of properties that are not derivable from more general patterns, it can still be considered as a particular instantiation of the resultative construction, because the latter does in fact constitute a rule of the language, and the way construction inherits its general characteristics.

It is true that there are other grammatical patterns in Dutch with which the \textit{weg}-construction shares properties, so that it may be said to be not completely isolated. For example, there are reflexive expressions similar to the way construction but lacking the way-constituent itself:

\begin{verbatim}
(8) Zij worstelde zich een weg door de menigte.
    She struggled \textit{refl} a way through the crowd
    'She struggled her way through the crowd.'

(9) Zij worstelde zich door de menigte.
    She struggled \textit{refl} through the crowd
    'She struggled through the crowd.'
\end{verbatim}

Still, the way construction present in (8) is not \textit{just} a special case of the 'reflexive movement' construction exemplified in (9), since the former implies exertion of energy and removal of obstacles, while the latter does not:

\begin{verbatim}
(10) Zij bewoog zich rustig naar de uitgang.
    She moved \textit{refl} quietly to the exit
    'She quietly moved to the exit.'
\end{verbatim}

Another related set of expressions are those in which a so-called 'possessive dative' occurs, and which may therefore suggest that the difference between Dutch and English is one of a more general pattern:

\begin{verbatim}
(11) Het hart klopte hem in de keel.
    The heart \textit{beat-past} him in the throat
    'His heart was in his throat.'
\end{verbatim}
After all, the role of the reflexive in the Dutch *way* construction may just as well be called possessive (indicating the one whose way is made; the fact that English uses a possessive construction proves the possibility of such a construal). However, while this is a productive construction in German, it is not so in Dutch, just as it is not productive in English; there are only fixed, idiomatic expressions, and no template (Hüning 2003:151–153). So although the Dutch construction certainly belongs to a number of larger families, it is still not a straightforward instantiation of a more general rule of the language.3

The *way* constructions in Dutch and English thus provide a nice illustration of the fact that the grammars of two languages may differ even though they share the rules as such (in a particular domain): the way the constructions are connected to the rest of the network differs. Clearly, a difference at a relatively general level, such as "English has a dative benefactive construction, Dutch has not" does not predict the existence or non-existence of productive patterns at a more specific level, as the comparison with Dutch shows. The theoretical consequence of this is that the distinction between being or not being an instantiation of general rules in a language – the place of a construction in ‘the system’ – is only a minor factor determining the status of a grammatical construction.

3. The *time-away*-construction

According to Jackendoff (1997), the *way* construction belongs to a set of constructions sharing several features of form and meaning. Another member of this set, he claims, is what he calls the *time-away* construction; examples are in (13) and (14).

(13) Bill slept the afternoon away.

(14) We’re twistin’ the night away.

3. One might suggest that the difference between English possessive and Dutch reflexive/benefactive marking also entails a difference in semantic content (construal). However, I see no empirical basis for such a claim at present. So for the time being, I consider this an illustration of what Croft and Cruse (2004:73) call the “conventionalist universalist position” Undoubtedly, “[w]hen a grammatical structure is used for the first time […], it does influence the way speakers think […]. But […] as it becomes the normal or even the only way to talk about the experience – then the original construal no longer constrains how speakers think of that experience”. To give another example: the fact that the conventional way of talking about ‘showing’ something in Dutch is to use a causative expression (*laten zien*) does not seem to me to imply that the Dutch generally have a more analytic kind of conceptualization of such events than speakers of English.
English constructions from a Dutch perspective

Figure 6. The time-away construction in the English constructional network

On the one hand Jackendoff claims a separate status for this construction, while on the other he maintains the position that it shares important properties with other patterns in the language:

This construction shares many general properties with the resultative construction [...] and the way-construction, [...] however, [...] it is a distinct member of a family of constructions to which all three belong. (Jackendoff 1997: 53)

We may depict the network suggested by Jackendoff as in Figure 6.

An independent status for the time-away construction is justified, as usual, by the fact that a specific set of characteristics of form is conventionally associated with a specific function. In this case, the template contributes the meaning that the object in question is restricted to indications of time periods, and that the subject referent wastes this period by engaging in the activity indicated by the verb. Thus, English has a syntactic template that (using the notation from Goldberg 1995) can be represented as in (15):

\[(15) \begin{bmatrix}
    \text{Sem: actor, waste, time, lost} \\
    \text{Syn: SUBJ V OBJ away}
\end{bmatrix}\]

Apparently, Dutch does not have an analogue to the time-away construction – at least not a syntactic one. But there does seem to be a morphological process playing the same role as the English syntactic template. One of the morphological categories analysed in De Vries’ (1975) book on Dutch verbal morphology, is exemplified in (16)–(17), as is clear from these examples, this could at least count as the translation equivalent of the time-away construction.

\[(16) \text{Hij had de hele middag verslapen.}
\text{He had the whole afternoon ver-slept}
\text{‘He had slept the entire afternoon away.’}\]

I would like to thank Ariane van Santen for drawing my attention to De Vries’ analysis.
They have their time *ver*-chattered

‘They chattered their time away.’

By prefixing a verb with *ver-*, the idea is evoked that what the object refers to is spent completely (in fact wasted) through the process the subject referent chose to engage in (denoted by the verb-stem). This raises the following question: How much does the fact that the English construction seems to belong to a family of syntactic patterns and the Dutch one to morphology, actually contribute to their identity? Does it help to explain anything about the differences between the two, beyond what is implied by what makes one a case of syntax and the other one of morphology? In discussing the syntactic difference between the Dutch and English way constructions, I noticed that they seemed to occupy rather different positions in the grammatical space of each language (one instantiating a ditransitive, the other a transitive pattern), but concluded that this actually only shows the limited relevance of abstract patterns in determining the properties of a grammatical construction. In the same vein, we may also doubt the importance of the general distinction between ‘syntax’ and ‘morphology’ for the nature of the Dutch and English constructions considered here. We could simply analyse the constructions in the two languages as in (15) and (18), and claim that this basically captures what the difference amounts to.

(18) The Dutch constructional template mentions parts of words, and specifies that *ver-* and the verb stem combine into a single word. This implies, among other things, that no other material may intervene between these two elements; more generally, in the absence of specifications to the contrary, the combination behaves in the way single words normally do in the language. But beyond that, no further consequences need to follow; the way this construction interacts and combines with other elements (words or templates) is not particularly different from the way the ‘syntactic’ English construction does. Constructions do not divide into two classes with different combinatorial or functional properties corresponding to the distinction between morphology and syntax. In a constructionist view, similar grammatical behaviour results from similarity of the symbolic relations involved rather than from the position of the construction in the grammar as a whole, so we should expect that the ways in which the English construction (15) and the Dutch one in (18) are used and combine with other elements are similar. In fact, they are. First, consider the verbs listed in (19).

---

Such low level specifications preventing inheritance of default properties of words may, for example, involve non-standard stress patterns, but also something like ‘separability’ of the ‘prefix’ (which is therefore usually referred to as a ‘particle’), which has syntactic consequences. For an analysis of Dutch ‘separable complex verbs’ as constructions, see Booij (2002), Section 3.3.
Table 1. Verbs with prefix ver- in Dutch and their objects

<table>
<thead>
<tr>
<th>Semantics of object</th>
<th>Semantics of verb (ver-V): &quot;waste (by V-ing)&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Chances, rights, reputation [± basis for future well-being]</td>
<td>verspelen (N.B. ( \neq ) ver+spelen, &quot;play, gamble&quot;!), &quot;waste, throw-away&quot;.</td>
</tr>
<tr>
<td>2 +Time/–Money</td>
<td>verdoen (met, &quot;with&quot; X), &quot;do (X)&quot;; verdromen, &quot;dream&quot;; verlummelen, &quot;hang around&quot;; verslapen, &quot;sleep&quot;.</td>
</tr>
<tr>
<td>3 +Time/(+Money)</td>
<td>verkletsen, verlullen, &quot;chatter&quot;; verpraten, &quot;talk&quot;.</td>
</tr>
<tr>
<td>4 –Time/+Money</td>
<td>verdobbelen, &quot;play dice&quot;; vergokken, &quot;gamble&quot;; verroken, &quot;smoke&quot;; verschieten, &quot;shoot&quot;; versnoepen, &quot;eat candy&quot;; verspelen, &quot;play, gamble&quot;; verzuipen, &quot;drink (alcohol)&quot;.</td>
</tr>
</tbody>
</table>

(19) verkwanselen, \"bargain/fritter away, waste\"; verkwisten, \"waste, squander\"; verspillen, \"waste, fritter away\".

These are members of the relevant category, but they cannot be considered as blends of the construction and other elements, simply because the latter do not exist. The verbs share an aspect of form corresponding to an aspect of meaning – the prefix ver- and the concept of wasting – but there are no verbs kwanselen, kwisten en spillen. Thus, the role of the prefix here is one of formal similarity indicating semantic similarity, not one of a compositional element of the verbs. As Jackendoff points out, there are verbs in English, such as while and fritter, that occur only in the time-away construction, and this is of course precisely the same kind of situation. People have specific instances of the constructions, including their structure, stored in memory, as well as (slightly) more general patterns, some of which may be used productively.

Next, consider the data in Table 1, which lists the kinds of objects that some typical instances of the Dutch construction (taken from a large dictionary, Van Dale) apply to.6

There are some specific instances of conventionalization here, such as the specific association of the verb verspelen with objects that, apparently, evoke some notion of ‘opportunities’ being wasted, without a specification of the nature of the activity that produces this result. However, the same form verspelen may also be combined with an object that does not refer to time, but rather represents some monetary value (e.g. “a week’s pay”), in which case the activity spelen, \"play\", is understood in its specific sense of “gamble”. There are several verbs taking this type of object (category 4 in Table 1). Also, verbs indicating some form of talking (category 3) normally take objects indicating a period of time, but they may also mention monetary values in the ‘right’ contexts, especially when time literally costs money: one can “chatter away” a fortune making long distance telephone calls.

6. Morphological categories have the advantage that it is relatively easy to use lexicographic and corpus tools to investigate them.
Beyond these specifics, however, there is clearly also a general pattern: the objects in instances of this construction indicate things that are considered *valuable*; (18) can be generalized to (20).

(20) \[
\text{Sem: actor, waste, valuable} \\
| cause | \\
\text{Syn: SUBJ [ver[V]] OBJ}}
\]

The fact that these often, and especially with certain kinds of activities, involve time, does not have to come as a surprise, given the metaphorical mapping between time and money. Now according to Jackendoff, the English construction, as indicated by the name he gives it, requires objects indicating *time*. Although it cannot be excluded that such a specific kind of meaning is conventionally tied to an expression while a conceptually related one is not, it is not the kind of thing to be expected, in the view presented here. Productive patterns express repeated conceptually salient aspects of situations, so it would be surprising if this shared cultural model – the time-as-money metaphor – would completely fail to show up in this pattern in English. And in fact, such examples do occur. A typical instance of the Dutch construction with its English translation is (21), and example (22) stems from the Brown corpus.

(21) Hij vergokte zijn erfenis.  
He *ver*-gambled his inheritance  
‘He gambled his inheritance away’

(22) But it is our health – more precious than all the money in the world – that these modern witch doctors with their fake therapeutic gadgets are gambling away.

Example (22) is especially interesting in that it not only shows that the ‘X away construction’ is available for expressing the idea that something valuable has been wasted because of someone’s activity, but explicitly invokes money, the prototypical instantiation of valuable commodities, as the standard of comparison. It seems clear that in both languages, other valuable things than time occur in instances of this pattern, so that the semantic specification of the construction represented in (15) must be revised as in (23):

(23) \[
\text{Sem: actor, waste, valuable, lost} \\
| cause | \\
\text{Syn: SUBJ V OBJ away}}
\]

Still, it is conceivable that speakers of English also have (15) stored as a specific instance of (23), possibly its prototype. Thus, there may be differences between the functions
of the Dutch and English constructions, but such differences would then concern the level in the taxonomy of constructions at which ‘time’ is a conventionalized aspect of meaning or its degree of entrenchment, not its presence as such.

At a relatively specific level of the constructional patterns themselves, we have once again seen a remarkable parallel between Dutch and English. In this case, the Dutch perspective even offered a possibility to improve the analysis of English. So let us return to the question about the role of higher level generalizations as possible factors determining the properties of the constructions. As we have seen at the beginning of this section, Jackendoff claims that there are many such generalizations, suggesting that this captures an important insight. But it is quite unclear what kind of consequences, i.e. predictions about differences in grammatical, combinatorial behaviour, follow from such a statement. In fact, Jackendoff is much more successful in pointing out the peculiarities of the construction than its similarity to others; claims of the type “Important properties such as X are explained by the fact that this construction is related to that one” are conspicuously absent from his treatment. Assigning an important role to the general patterns would seem to entail the prediction that when such general patterns are different in another language, the constructions should also differ in other forms of grammatical behaviour, and that simply seems hard to substantiate.

4. Causative constructions

In the case of English, the two sets of specific constructions discussed in the previous sections appear to be well integrated into the system of the language. Even though the general rules may not determine the grammatical properties of the specific patterns to a large extent, the latter do instantiate the former: the way and the time-away constructions are both cases of the transitive (resultative) pattern. Their Dutch counterparts appear to be less well integrated. Although it cannot be excluded in principle, it is unlikely that this asymmetry is a consequence of general properties of the languages (English allegedly being ‘more systematic’ than Dutch). Let me briefly mention one example in which the situation seems to be reversed, viz. the case of causative constructions: clauses in which a basic verb form – stem or infinitive – is combined with a marker (a special kind of verb, or a specific affix) to yield an expression indicating the causation, in one way or another, of the process or state indicated by the verb stem or infinitive (which thus functions as a ‘result predicate’). Some examples in English are *He made me do it*, *She let me sleep for an hour*, *He had Da Vinci paint his portrait*.

It might be, for example, that the concept of “wasting” is slightly more prominent in the Dutch construction than in its English counterpart. Cf. Jackendoff’s (1997:537) characterization “the subject is in some sense understood as […] ’using the time up’. Some of this flavour appears in the second-approximation paraphrases […] Sam spent/wasted the afternoon sleeping.”
Kemmer and Verhagen (1994) explore the way in which the roles of participants in causative constructions are marked. In this respect, causative expressions quite generally turn out to share many properties with simplex clauses, not with complex clauses, contrary to what is predicted by many analyses assuming an underlying dual clause structure for causatives. Causes are frequently marked by means of case endings or prepositions that are available in a language for other purposes, in ways that are motivated by the meanings such markings have in simplex clauses. In particular, markings for participants with a somewhat ‘intermediary’ role in the causal structure of events — roughly, ‘datives’ and ‘instrumentals’ — are often also employed to mark causees; (24) and (25) give examples of ‘dative-like’ marking in French and Hindi respectively:

(24) J'ai fait manger les pommes à Jean.
   I have made eat-INF the apples to John
   ‘I made John eat the apples.’

   I AG Ram-DAT spice taste-CAUS-PAST
   ‘I had Ram taste the seasoning.’

Examples (26)–(28) illustrate the use of ‘instrument-like’ markings in German, French, and Hindi, respectively:

(26) Er liess den Brief von seinem Sohn abtippen.
   He let the-ACC letter from his-DAT son type-INF
   ‘He had the letter typed by his son.’

(27) J'ai fait manger les pommes par Jean.
   I have made eat-INF the apples by John

(28) Mai-nee raam-see masaalaa cakh-vaa-yaa.
   I AG Ram-INST spice taste-CAUS-PAST
   ‘I had the seasoning tasted by Ram.’

Table 2 (from Kemmer & Verhagen 1994:137) gives an overview of causee markings (for nominals – pronouns often exhibit different marking possibilities) in a number of different languages.

The table only indicates whether a marking can occur at all in causative constructs. A ‘+’ for ‘dative’, for example, does not mean that this marking is possible in all cases, nor that it occurs in exactly the same range of cases as in another language. Full predictability does not exist, as the examples given already illustrate, and witness the case of German: this language does have a separate dative case but employs accusative for causes, so that causatives of transitive verbs exhibit double accusative marking. Nevertheless, there is clearly a motivated pattern. In terms of the present approach, causative constructions in many languages are specific instantiations of simplex clause constructions, with some specific properties of their own and otherwise inheriting general properties of their simplex clause parents, especially those subtypes that are semantically similar. That is, what we can learn from Table 2 (along with
English constructions from a Dutch perspective 271

Table 2. Marking of causees in causative constructions in a number of languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Ø</th>
<th>'ACC'</th>
<th>'DAT'</th>
<th>'INST'</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>German</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Hindi</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Kannada</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Mongolian</td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Quechua</td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Dutch</td>
<td>+</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>English</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

other evidence discussed in Kemmer & Verhagen 1994), is how causative con-
structions are normally integrated (in the sense mentioned before) into the grammar of a
language: they tend to share properties with two types of general simplex clause pat-
terns, viz. indirect object constructions and instrument-like constructions available in
the language involved.

Now, the situation in Dutch looks rather similar to that in most other languages
listed in Table 2.8 The specific indirect object marking with the preposition _aan_ ('at',
'to'), for example, can occur in causative sentences if the causee's role in the overall
process is sufficiently similar to that of a recipient, also when none of the verbs involved
by itself specifies such a role, as in example (29).

(29) Je mag die brief _aan_ niemand laten lezen.
     'You may not let anyone read that letter.'

But as Table 2 indicates, English causative constructions do not inherit this kind of
property from general (simplex clause) recipient constructions. This indicates that
English causative constructions are more like an island of constructions in the entire
constructicon than the Dutch ones. The English construction does not share as many
properties with productive basic clause patterns of the language as the Dutch one does.

Another phenomenon strengthening this conclusion concerns so-called causeeless
causatives. Many languages allow the causee to remain implicit in the case of transitive
result predicates, and Dutch is no exception, witness example (30):9

(30) We laten een nieuw huis _bouwen._
     'We are having a new house built.'

9. Dutch _laten_ has a wider meaning than English _let_, and also subsumes coercive readings. Cf.
The optionality of the causee reflects the optionality of recipient participants, or indirect objects in general, in transitive clauses, i.e. clauses containing a direct object. It is thus a further indication that the Dutch causative construction is relatively well integrated in the general network of simplex clause constructions in Dutch. English, on the other hand, is again an exception in this regard, in that it does not allow causees to remain implicit in standard causative constructions. This does not imply that a causative event could not be described in English without mentioning the causee; but one has to use another construction to do this: as the translation in (30) indicates, the English counterpart to the Dutch causeless causative contains a past participle rather than an infinitive/stem. Thus, this constitutes a separate member of the English network of causative constructions, setting this family as a whole somewhat more apart from other simplex clause constructions than in a language like Dutch. The use of the past participle, combined with the fact that the causee remains implicit, constitutes a resemblance between this subtype of English causatives and passive constructions, linking this particular family of causative constructions to another part of the constructicon than the causatives based on infinitives/stems. The fact that the causee may be inserted in this construction in a by-phrase (e.g. He had his portrait painted by Da Vinci, and the translations of (26) and (28)), strengthens this link.

Naturally, this is not an exhaustive discussion of causative constructions in English, but it suffices to see that in this area, Dutch may be considered more ‘systematic’ than English. More important, though, is the fact that lack of full-scale regularity is, apparently, not really a problem in either system.

5. Conclusion

Construction grammarians have discovered many phenomena of idiomaticity at practically all levels of schematicity in grammars. They have also recognized the theoretical importance of these phenomena, and of the fact that they are widespread (e.g. Jackendoff 1995; Langacker 1988). There is now considerable consensus on the crucial role of constructions in the general organization of grammars, and in the processes by which grammars are acquired and change over historical time. But there also remains a lot to be investigated. In a sense, the constructionist approach has extended the number of interesting phenomena to be studied, because of the insight that grammars do not only consist of regularities on the one hand, and idiosyncracies on the other. Rather, some combination of the two seems to be the rule rather than the exception (paradoxically so), so that the balance is always an open issue, and thus deserves investigation. In this paper, I have shown that a comparative perspective can help us get a better understanding of this balance in a number of grammatical domains. Some constructions of English are more tightly integrated into the constructicon than others; the same holds for Dutch. But the most regular constructions in one language certainly do not have to be the same as those in another. Thus, a complete grammatical system
may differ considerably from another one even though they may both contain largely the same set of patterns.

From the authors cited in this paper, Jackendoff most frequently invokes similarities of constructions to other patterns in English, at least suggesting that these are important factors in understanding the construction involved. However, we have seen that at relatively specific levels, constructions share important combinatorial and functional properties across languages, whether they are syntactically similar or not. Different positions of constructions in the overall taxonomic network, and different degrees of integration into this network, turn out not to provide clear and consistent clues about the grammatical or semantic behaviour of the constructions – the role of convention is bigger than apparent formal systematics within a single language may suggest. From a cognitive point of view, emphasizing the symbolic character of constructions, this need not come as a surprise. Even if speakers of different languages have to use different conventional resources, if the kind of concepts they regularly express – e.g. actions of obstacle-removal, wasting valuable resources, making things happen or having people do things – show systematic (cross-cultural) similarities, then the behaviour of these different resources can be expected to exhibit similarities.

In principle, constructional grammars are well designed to handle the kind of phenomena dealt with here. Similarities are captured by inheritance relations between more general and more specific patterns. Such relations will occur frequently in the grammar of any fully developed language, because of the way the language has developed over time (conventionalization of routines that started as fully 'compositional' expressions; cf. Verhagen 2003:53–56), and because of the usefulness of motivation: a system is more efficient if formal distinctions correspond to semantic ones, i.e. if the formal distinctions are motivated (Haiman 1985; Lakoff 1987:537–540). There are also other mechanisms that provide 'pressure' for regularity, especially the way children construct knowledge of grammatical structure (Tomasello 2003).10 So there is bound to be a lot of regularity in any grammar, rather than complete arbitrariness of the form-meaning-relationships in each separately stored construction. However, the constructional approach does not impose a priori constraints on which aspects of these relationships will be shared by many constructions at different levels of schematicity and productivity, which is why it is well suited to deal with cross-linguistic differences of the kind I have illustrated here.

References


10. See Goldberg (1995:Ch. 3) for an overview of relevant factors, as well as Goldberg (2006).


Goldberg, A. E. (1996). Making one’s way through the data. In M. Shibatani & S. A. Thompson (Eds.), Grammatical Constructions: Their Form and Meaning (pp. 29–53). Oxford: OUP.


Notes towards an incremental implementation of the Role and Reference Grammar semantics-to-syntax linking algorithm for English*

Christopher S. Butler
Centre for Applied Language Studies, University of Wales Swansea, UK

This paper presents some ideas which might be incorporated into a functional model of language production based on an incremental application of the semantics-to-syntax linking procedure of Role and Reference Grammar (RRG). After a brief outline of syntax, semantics and their interrelationship as seen in RRG, the idea of incrementality in language production is introduced, with particular reference to the work of Ferreira (2000). A dynamic implementation of the linking procedure for a simple sentence is then proposed. The problems raised by information (topic, focus) structure in more complex sentences are then addressed, and an example is worked through in detail. The paper concludes with a brief section which sets this work within the context of a wider research programme.

1. Introduction

In Butler (2003a:487–488) it is argued that if, as proposed in Dik's Functional Grammar (FG) (see Dik 1997a:13–14) and Van Valin's Role and Reference Grammar (RRG) (see Van Valin & LaPolla 1997:13), a structural-functional model of language is to

* The work reported in this paper formed part of a larger project, BFF2002-02441, Análisis del discurso en lengua inglesa: aspectos cognitivos, contrastivos y de adquisición (Analysis of discourse in English: cognitive, contrastive and acquisitional aspects), which was financed by the Spanish Ministry of Education, with further support from the Xunta de Galicia (XUGA, grant reference PGIDIT03PXIC20403PN). I wish to acknowledge not only the support provided by these authorities, but also my indebtedness to Lachlan Mackenzie, the dedicatee of this paper, and to an anonymous referee, for their valuable comments on an earlier draft.
strive for psychological/cognitive adequacy, it must take fully into account the evidence on language processing which is available in the psycholinguistic literature. This evidence suggests that the grammatical model should contain levels of semantic, morphosyntactic and phonological patterning, as well as a lexicon which integrates information from all other linguistic levels and has pointers from lexical senses to bundles of concepts.

It is further argued (Butler 2003a:497–498) that existing structural-functional models fail, in various ways, to provide all the required components. Systemic Functional Grammar (Halliday & Matthiessen 1999, 2004; Matthiessen 1995) has no lexicon as such, treating lexical items as realizations of clusters of delicate selections from system networks which also embody grammatically-realized distinctions. Mainstream FG (Dik 1997a, 1997b) has no separate level of syntactic patterning, relegateing syntax to the insertion of the syntactic functions Subject and Object in the underlying clause structure, and the operation of the rather poorly worked out expression rule component.¹ RRG (Van Valin & LaPolla 1997; Van Valin 2005) does indeed have the required components, but suffers from a serious lack of detail in its specification of lexical items. Given the important role which the lexicon is generally agreed to play in language processing, this is an important drawback. Particularly interesting, then, is the model proposed by Mairal Usón and his colleagues (Mairal Usón & Van Valin 2001; Mairal Usón & Faber 2002, 2005; Mairal Usón 2004), which synthesizes the rich lexical specifications of the Functional Lexematic Model (FLM, itself a hybrid of FG and Coseriu’s Lexematics: see Martín Mingorance 1998; Faber & Mairal Usón 1999) with the semantic and syntactic apparatus of RRG, including cross-linguistically tested linking principles for going from semantic representations to syntactic representations and vice versa. The question is whether these principles lend themselves to a dynamic interpretation which provides an appropriate model for the production and comprehension of language. Van Valin (2006) argues that there is a considerable degree of correlation between the stages in the RRG semantics-to-syntax linking procedure and the two levels of grammatical encoding (functional processing, positional encoding) proposed in the production model of Levelt (1989). The problem with such a simple

¹. Other scholars working within the FG framework have, however, made a number of proposals aimed at increasing the level of psychological adequacy. The Functional Procedural Grammar of Nuys (1992, 1994, 2001) aims to model the cognitive systems involved in language production. The Functional Discourse Grammar of Hengeveld (Hengeveld 2004a, 2004b, 2005; Mackenzie & Gómez-González 2005; Hengeveld & Mackenzie 2006, in preparation), though itself purely a model of the discourse grammar, envisages the possibility of an implementation of the model which would mimic production. The Incremental Functional Grammar of Mackenzie (2000, 2004) aims to increase psychological adequacy by recognizing the incremental nature of language production. The model of the natural language producer proposed by Bakker and Siewierska (2004) also involves incremental expression of the underlying structure of a clause. A comparison of these proposals with those made here would take us far beyond the scope of the present article, and is left for future research.
correlational model, however, is that it does not take account of the fact that production is generally taken to be incremental, in the sense that partial information is sent from one level of processing to the next, before all the information at a particular level is available. The aim of the present paper is to suggest how the RRG linking algorithm might be implemented in such a way that the principle of incrementality is respected.

In Section 2, I present a brief outline of syntax, semantics and their interrelationships in RRG. Section 3 then refers briefly to the issue of incrementality in language processing. Section 4 then shows how we may implement the RRG linking algorithm dynamically in order to model certain aspects of language production, with particular reference to English. In Section 5 I discuss some problematic issues and suggest ways in which we might begin to deal with them. In the light of the discussion in this section, Section 6 then presents a rather more complex example of the dynamic implementation of the linking algorithm. Section 7 concludes the paper and situates the work within the context of a broader research project.

It should be emphasised that what is presented here is not intended as in any sense a full, formalised model of the conceptual and grammatical aspects of production. Clearly such a model would require many additional features: for instance, Fortescue (this volume) demonstrates that top-down mechanisms are required to supplement the bottom-up building of structure, and indeed it is quite possible that a model whose grammatical component is based on RRG could accommodate, as a phenomenological basis, some of the Whiteheadian ideas put forward in Fortescue’s work. The goal of this paper is merely to offer a sketch of how the concept of incrementality in processing can be reconciled with an RRG-based approach to the grammatical aspects of production.

2. Semantics, morphosyntax and the lexicon in RRG

As we have seen, RRG has separate levels of semantic and syntactic representation. The semantics of a clause is represented by the Logical Structure (LS) for that clause, consisting of lexical representations for predicates and their arguments, and values for operators. In (2) below is the LS for the clause in example (1):

(1) They were walking in the street.

(2) \langle IF DEC \langle TNS PAST \langle ASP PROG (be-in’ (street, \langle do’ (3pl, \langle walk’ (3pl)\rangle)\)\)\)\)\)\)

This shows us that we have a predicate walk’ with a single argument (the 3rd person plural pronoun they), the whole of this structure being embedded within the general activity predicate do’, and that the adjunct in the street has scope over the whole of this. The values of the illocutionary force, tense and aspect operators are given as Declarative, Past and Progressive respectively. It is important to realize that elements such as walk’ and street are not intended to represent the actual English words walk and street, but are placeholders for a detailed lexical decomposition in terms of universal primitives, which, as we saw earlier, is currently being worked on by Mairal Usón and his
colleagues. This represents a major difference between RRG and mainstream FG, in that the latter operates with the principle of stepwise lexical decomposition, according to which the meanings of complex lexemes of a language are analysed in terms of the meanings of other, simpler lexemes (see Dik 1997a:97–103).

At the syntactic level, the clause is claimed to have a layered structure. The primary split is between the core and an optional periphery.\(^2\) The core consists of the nucleus, which houses the main predicate, and of one or more core arguments. For some languages, there is also a pre- and/or post-core slot within the clause; English, for example, has a pre-core slot into which go wh-items and fronted NPs and PPs. Again for some languages but not others, there may be left- and right-detached positions which precede or follow the clause proper. Van Valin presents arguments, from a number of languages, in support of these distinctions (see Van Valin 1990; Van Valin & LaPolla 1997:31–40; Van Valin 2005:6–7, and the brief summary in Butler 2003b:123–128, 206–209).

The syntactic constituent projection\(^3\) for the clause in (1) is shown in (3):

![Syntax Diagram](image)

They were walking \(\text{in} \) the street

Syntactic structures for a language are stored in the syntactic part of a set of constructional templates, analogous to the constructions of Construction Grammar (see e.g. Goldberg 1995, 1996), which can also contain semantic information (meaning of the

---

2. In the latest version of the theory, each syntactic layer of the clause (clause, core, nucleus) can have a peripheral element attached (Van Valin 2005:21). I shall not pursue this idea further here.

3. It is important to realize that the constituent projection of RRG does not represent an ‘immediate constituent’ analysis, which Van Valin & LaPolla (1997:23–25) criticize on the grounds of the serious problems it brings. Rather, it refers to the constituents of the layered structure of the clause as outlined here.
construction, any semantic constraints on it), pragmatic properties (related to illocution or focus) and any morphological information peculiar to the construction. The syntactic component of the templates reflects not only the universal layered structure of clauses (core and periphery, nucleus and core arguments within the core), but also any language-specific components such as left- and right-detached elements. For English, the templates would include those for left- and right-detached positions, the pre-core slot and the various possibilities for the structure of the core. The structure in (3) is based on the Core-4 template from the set given by Van Valin (2005:15). It has just one argument position, and this is in accordance with the following syntactic template selection principle:

The number of syntactic slots for arguments and argument-adjuncts within the core is equal to the number of distinct specified argument positions in the semantic representation of the core. (Van Valin 2005: 130)

Note that the principle refers to the number of distinct specified positions in the LS: the LS in (2) has arguments for both do´ and walk´, but these are the same, namely '3pl', so there is only one distinct argument.

It will be noted that the auxiliary were is not attached to the constituent projection for the clause. Such elements arise as realizations of operators, shown in a separate operator projection which complements the constituent projection. For the argumentation behind this proposal, which relies heavily on cross-linguistic generalizations, see Van Valin & LaPolla (1997:70–71) and Van Valin (2005:11–13). In addition to the constituent and operator projections, RRG postulates a third projection which shows the potential focus domain (PFD) which the language allows, and the actual focus domain (AFD) for a particular linguistic example. For English, the potential focus domain is the whole clause, whereas for Italian, for instance, it is the verb plus any postverbal material.

The complete structure for example (1), with constituent, operator and focus projections, is shown in (4) below: it is assumed for purposes of exposition that the actual focus domain is in the street, though there are of course other possibilities. The operators for progressive aspect, past tense and declarative illocutionary force are shown as attached to particular levels within the layered structure.
(4)

Noun phrases also have both a semantic structure and a layered constituent projection and operator projection, for which arguments can be found in Van Valin & LaPolla (1997:53–67) and Van Valin (2005:24–30). The layered structure for the simple NP *the street* in (1) is shown in (5), and its semantic structure in (6) (see the similar example in Van Valin 2005:52). The operators in the semantic structure are for definiteness (value positive), negation (zero), quantification (existential), number (singular) and nominal aspect (count).
It is important to note that the logical structures for predicates, whether verbal, nominal or adjectival, form the heart of the entries for these predicates in the lexicon.

The semantic and syntactic representations for a sentence are related through two sets of linking procedures, one going from semantics to syntax (corresponding to language production), the other from syntax to semantics (corresponding to comprehension). In order to understand how these linking procedures work, we need to introduce the concepts of 'thematic relation', 'macrorole' and 'privileged syntactic argument'.

Thematic relations such as Agent, Patient or Theme are not seen in RRG as primitives of the theory, since they are claimed to be derivable from the configurations of the LS, which depend on the lexical class of the predicate, which can in turn be decided according to a set of rules. Let us see how this works for example (1). The predicate walk in (1) is classified as an activity in RRG, based on the combined results of a battery of tests for Aktionsart: it can occur with the Progressive, as in our example; it also co-occurs with adverbs such as vigorously and with pace adverbs such as slowly; it can take expressions such as for an hour but not in an hour. Activity predicates all have an LS of the type do’ (x, [predicate’ (x) OR (x, y)]), depending on whether they are 1-place or 2-place. The first argument of such a structure (the only argument in the case of a 1-place predicate, as in our example) always has the thematic relation of EFFECTOR, which can be subclassified according to the type of activity: in the present case, the argument they is subclassified as MOVER. We can go through a similar process for any predicate in any clause. Thematic relations such as EFFECTOR, PATIENT, THEME, etc. can be grouped under two ‘macroroles’, one representing a generalized AGENT-
Christopher S. Butler

Figure 1. The Actor-Undergoer Hierarchy (from Van Valin 2005:61)

![Diagram of Actor-Undergoer Hierarchy]

Type role labelled as Actor, the other representing a generalized PATIENT-type role labelled as Undergoer. The allocation of Actor and Undergoer status to arguments in a LS is performed according to the Actor-Undergoer Hierarchy, shown in Figure 1.4

From this, we see that the first argument of an activity predicate, as in example (1), will be the Actor. If there had been a second argument, this would have been the Undergoer. The categories of Actor and Undergoer are extremely important in RRG because they figure in a large number of explanatory statements.

We turn now to syntactic relations (see Van Valin 2005:Chapter 4). Proponents of RRG claim that syntactic relations should be postulated for a language only when there is at least one case of restricted neutralization of two or more semantic roles for syntactic purposes. Such is the case for English: for example the finite verb clause agrees in number and person with the NP at the beginning of the syntactic core, whether this is the Actor or the Undergoer, and this neutralization is restricted, since it applies only to arguments acting as one of these two macroroles, and not to any other arguments of the predicate. Some languages, such as the Austronesian language Acehnese, do not have such restricted neutralizations and so are argued to have no syntactic relations. Even where grammatical relations can be recognized, RRG does not make use of the traditional categories of Subject and Object, since it is argued that these cannot be justified cross-linguistically. Rather, RRG postulates a category of ‘privileged syntactic argument’ (PSA),5 tied to particular constructions in the language. In some languages, including English, most constructions have the same PSA (hence the conflation of the PSAs into the category of Subject), but there are languages which have a number of PSA types for different constructions.

Armed with information about macroroles and PSAs, we may now see how the linking algorithms work for our example.6 Irrespective of the direction of mapping, the first step is to determine the semantic representation for the sentence and the appropriate syntactic template.

4. In this hierarchy, DO is a special predicate used for cases where an agentive interpretation is obligatory, such as murdering someone.

5. Privileged syntactic arguments can be of two main types, pivots and controllers, but this distinction is not important for the points being made here.

6. The linking algorithms are, of course, much more complex than is suggested by the simplified account given here, which concentrates on just those aspects relevant to the example under discussion. For full details see Van Valin (2005:Chapter 5).
Let us now look at the semantics-to-syntax linking procedure for our example. Starting from the LS in (2), we determine the Actor (and if necessary Undergoer) assignment by consulting the Actor-Undergoer Hierarchy: we have seen that where there is an activity predicate with a single argument in the semantic structure, this argument will be the Actor. We now select the morphosyntactic coding of the arguments: the default in English is for the Actor to be the PSA, and this is in fact the only possibility here, since there is only one argument in the structure. The core arguments of the predicate are now given their correct case markers or adpositions, according to the rules of the language: in the present case, this means giving nominative case to the PSA pronoun (i.e. *they* rather than *them*). Agreement marking is assigned to the auxiliary, realised as *were*. We now choose the appropriate syntactic template according to the template selection principle stated earlier: as we have seen, this results in the selection of the Core-4 template. The core argument is assigned its proper position, and (by default) we assign to the periphery the LSs for any predicates other than the main verb, in this case *be-in*. The skeleton of the linking pattern is shown in (7), in which the operators in the LS have been omitted for clarity.

\[ (7) \]

The syntax-to-semantics linking procedure starts with the syntactic structure in (3). Since there is only one NP in the core, this will be the sole macrorole. We retrieve from the lexicon the LS for the main predicate, i.e. *do`*(x, [walk`*(x)*])). The Actor-Undergoer Hierarchy tells us that the first argument of such a structure will be Actor rather than
Undergoer, as would be the case in an example such as *John died*. We also retrieve the LS of the predicative PP, i.e. *be-in*´(x, y), and assign the LS of the core as the second argument of this predicate, and the object of the preposition (i.e. *the street*) as the first argument. The linking is shown in (8).

Let us now summarize what has been said about RRG, in relation to the non-conceptual components we require for our model of the grammar. RRG has clearly separated semantic and syntactic representations, linked by algorithms which have been worked out on the basis of data from a large range of typologically diverse languages. It also allocates a crucial role to the lexicon, since the heart of the lexical entry for a predicate is the LS for that predicate, which, as we have seen, is the basis of the semantic structures with which syntactic representations can be linked. One weak point in the RRG model is clearly the lack of a well-developed lexical decomposition system using putatively universal primitives. This drawback is, however, currently being remedied by the proposals of Mairal Usón and his colleagues, which combine the richness of the Functional Lexematic Model with the semantic and syntactic apparatus of RRG.

A further problem, however, is concerned with the psychological/cognitive adequacy of the model, which requires us to take into account that language production and understanding are dynamic processes occurring in real time. Bakker and Siewierska, who have presented a model of the speaker which includes an incremental, left-
Incremental implementation of RRG algorithm

It is a widely held assumption in the literature on language production mechanisms that processing is incremental: that is, small chunks of a potential utterance, rather than the material for the whole utterance, are sent from one processing level to the next (see for instance the influential model of Levelt 1989, 1999). There is evidence from psycholinguistic experimentation, reviewed by Ferreira (2000), to suggest that the syntactic form selected by a speaker (e.g. active vs. passive) is influenced by the degree of availability of concepts in the conceptual structure. This in turn depends on the type of concept and also on its current role in the discourse, if any: concepts relating to prototypical, concrete, and especially animate entities are generally more highly activated than others, and of course concepts which have already been introduced into the discourse and are still topical will be particularly active. The higher the degree of activation, the more likely a concept is to be realized in Subject position in the clause: if the concept relates to the agent in some action, we get an active clause, while if it relates to the patient, we get a passive.

However, there is also evidence, again reviewed by Ferreira (2000), that articulation of an utterance cannot begin until at least the verbal lexeme has been retrieved: the production of the Subject of a transitive verb requires the verb to have been activated, but not necessarily its Object, a model which Ferreira dubs ‘moderate incrementality’. Yet further evidence suggests that in conditions where experimental subjects are not under pressure to produce utterances quickly, the planning unit may be even longer, amounting to a whole clause. Ferreira opts in the end for a model in which the most available concept is strongly attracted to Subject position, but in which the verbal concept must be activated, together with the corresponding syntactic argument structure affecting the whole clause, before articulation begins. Ferreira’s model is based on
Tree-Adjoining Grammar (see Joshi 1985; Frank 1992), which can be used to formalize a variety of different linguistic theories. It is interesting to note that Van Valin & LaPolla (1997:654) observe that Tree-Adjoining Grammar has been used to formalize the combination of elementary syntactic templates into more complex templates, and this is indeed exactly what Ferreira suggests in her article, though she adopts a Principles and Parameters/Minimalist framework rather than that of RRG. In the next section I develop an implementation of the RRG linking algorithms along the lines of Ferreira’s approach.

4. A dynamic implementation of the RRG linking algorithms

As an initial example of how we might achieve a dynamic implementation of the RRG linking algorithms, I shall take the very simple clause shown in (9) below, which has exactly the same syntactic structure as the example discussed by Ferreira (2000), *The dog bit a flower*, but has been chosen in order to illustrate some features of the lexical decomposition system of RRG.

(9) The dog ate a bone.

In the analysis of her parallel example, Ferreira starts from what she calls a ‘propositional representation’ for the clause. For our example, Ferreira’s analysis would be as shown in (10).

(10) event: eat (def/1/agent/topic: dog; indef/1/patient: bone; past)

Items such as eat, dog and bone are said to represent concepts, and while the status of the other categories is not discussed, their labelling would suggest that they are conceived as semantic/pragmatic in nature. If this is so, then Ferreira is presumably operating under the assumption that we do not need separate conceptual and semantic representations, a position endorsed in, for example, the Natural Semantic Metalanguage model of Wierzbicka and her colleagues (Wierzbicka 1988 and thereafter; Goddard & Wierzbicka 1994, 1997, 2002) and also in the work of Dik (e.g. 1990:234). There is, however, convincing evidence that two different levels of representation are indeed required: the scope of the present article does not permit further discussion of this important issue, but relevant arguments can be found in, for example, Nuyts (1990, 1992:223–236) and Levinson (1997). I therefore propose a purely conceptual representation at this stage, I have also included more detail of the degree of activation of the entity concepts, using the nomenclature of LaPolla (1995) employed by Van Valin & LaPolla (1997:200–201; see also Van Valin 2005:79), which in turn derives from the work of Lambrecht (1994), and through him, goes back to even earlier
proposals by Chafe and Prince. Furthermore, I have included an important component not present in Ferreira’s structure, relating to the illocutionary point of the utterance.7

(11)  \text{assert(eat (active (topical) /single/definite/affecter/: dog; brand-new unanchored/single/indefinite/affected/focused: bone; past))}

This representation is meant to convey that the speaker wishes to make an assertion, and that the content of the assertion is that a single, definite dog, serving as the Affecter in the process, and active in the discourse through already being topical, performed in the past an action of eating, and that the entity affected by this action is a single, indefinite bone, which is a brand-new referent unanchored to any previous referent, and is to be made the focus of the assertion.

According to Ferreira’s model the first concept to be processed will usually be the one with the highest activation, i.e. the topical one, dog. In RRG terms, this would mean that we retrieve from the lexicon the entry for the lexeme \text{dog}.8 The fact that \text{dog} is a noun means that it will be the head of a NP. We therefore invoke the syntactic part of the constructional template for a NP with lexical head:9

(12) NP \quad \text{CORE}_N \quad \text{NP} \quad \text{ADV} \quad \text{NP} \quad \text{PP} \quad \text{PP} \quad \text{NP} \quad \text{N}

The activation of the NP template brings with it the possibilities offered by the operators which work at the NP level: nominal aspect, number (sing/pl), quantification,

7. Inclusion of illocutionary point in this representation begs the important question of whether illocutions, attitudes and the like are actually conceptual in nature. In Butler (forthcoming) it is argued that they are not, but that the ‘content’ of what we say should be split into two types: conceptual content, relating to objects, events and properties in the world under description; and affective/interactional content, covering speech acts, attitudes, emotions, etc. I shall not pursue this distinction further here.

8. The lexical entries for nominal lexemes actually include a rich qualia structure of the type proposed by Pustejovsky (1991, 1995), but as this is not relevant to our concerns here it will be omitted.

9. In the NP template NPIP represents the ‘noun phrase initial position’ in which possessives and demonstratives are claimed to occur in English.
negation, definiteness and deicticity. There is also the possibility of one or more modifying adjectives and/or nouns, seen as modifiers in the periphery of the nominal nucleus (Van Valin 2005:26). The NP structure plus its operators thus allows the expression of a rich variety of modifications to the conceptual base of the NP, but our speaker wishes only to express the basic concept dog plus the positive value of the definiteness operator. The layered structure introduced at this point is thus as in (13):

(13) NP
    |   CORE_N
    |   NUC_N
    |   N
    the  dog

This NP will also be marked as Topic, in view of the feature TOPICAL in the conceptual specification. This structure is sent, not to the phonological mechanism, but to a buffer, while the verb is retrieved. The concept eat constrains the grammatical encoder to select from the lexicon the appropriate entry for the verb eat. There will in fact be two such entries in the lexicon, since eat has a non-telic sense in which only the activity of eating is involved (as in Pete ate pizza for ten minutes) and a telic sense in which the act of eating comes to completion, the thing which is eaten being wholly consumed in the process (e.g. Pete ate a pizza in ten minutes). The two senses have different logical structures. Since the clause in (9) is intended to have the second meaning, the encoder retrieves the following LS from the lexicon.

(14)  do’ (x, [eat’ (x, y)]) & INGR consumed’ (y)

This LS tells us that ‘x’ eating ‘y’ (itself an activity) results in a change of state whereby ‘y’ is totally consumed, and represents what is called an ‘active accomplishment’.
The semantics-to-syntax linking rules now come into play. We determine the Actor and Undergoer by means of the Actor-Undergoer Hierarchy, which specifies that the highest ranking argument in the hierarchy (here, the first argument of the do’ predicate, x) is the Actor, the other argument (y) the Undergoer. As we have an active clause, and English is an accusative rather than an ergative language, we get the default assignment of privileged syntactic argument to the Actor, that is: x = Actor = PSA. We also know that the value of the TENSE operator must be ‘past’.

It should be noted at this point that the lexical decomposition system of RRG has two features of relevance to the accuracy and efficiency of processing. Firstly, as we have just seen, it ensures the correct selection of the macroroles Actor and Undergoer, and ultimately correct syntactic function assignment (though see also the complexities introduced in §6). Secondly, it makes generalisations over sets of predicates (e.g. all activities have a similar decomposition), so enabling the rules to be stored and used in a more general form than would be the case if they operated on individual, non-decomposed verbs.

The illocutionary force operator will have the value ‘declarative’ in response to assert in the conceptual representation, and this means, for English, that the PSA argument will precede the verbal nucleus of the core. Furthermore, we know that the first argument of a do’ predicate will be an effector, the second argument of an activity predicate a patient.

We can now invoke a core template for the clause in order to accommodate the information we have so far. The required template to fit the data we have is Core-3 (Van Valin 2005: 15):

(15) CORE
    NP_NUC_NP
    PRED
    V

The core, in turn, fits within the template showing possible elements for the whole clause (16), which in turn fits into the template for the whole sentence (17) with which it is coextensive (in our sentence there is no element in the left-detached position, and neither is there any element in the pre-core slot of the clause):
So we have the following layered structure, including a compact representation of the operator projection (without the internal scoping relations for particular syntactic units), for a sentence consisting of just a clause headed by the active accomplishment predicate *eat*:

The PSA must be linked to the first core argument in an active clause. So far, then, we have the structure and linkages shown in (19):
(19)

Incremental implementation of RRG algorithm

```
the dog ate

Actor = EFFECTOR = PSA
Undergoer = PATIENT

do (x, [(eat (x, y)]) & INGR consumed (y)
```
The part of the syntactic structure so far established might go to the phonological system at this point, especially if the speaker is under pressure to produce the utterance quickly.

The structure so far predicts the need for a second entity which will fill the roles of Undergoer/PATIENT. The concept bone activates the lexical entry for bone. As for the dog, the NP constructional template is invoked, but here the semantic operator value -definite is assigned. The result is the following structure:

\[
(20) \quad \begin{array}{ll}
\text{NP} & \\
\downarrow & \\
\text{CORE}_N & \\
\downarrow & \\
\text{NUC}_N & \\
\downarrow & \\
\text{N} & \\
\downarrow & \\
\text{a} & \text{bone} & \\
\downarrow & \\
\text{N} & \\
\downarrow & \\
\text{NUC}_N & \\
\downarrow & \\
\text{CORE}_N & \\
\end{array}
\]

This NP can now be slotted into the free NP position in the core template for the whole clause, giving the final structure for the clause as shown in (21):
(21) The dog ate a bone.
We must also assign a focus structure consistent with both the communicative intentions of the speaker, and the rules of English. We have assumed that the dog concept is topical and the bone concept in focus. This is achieved by selecting, within the potential focus domain for English (the whole clause) an actual focus domain which extends over just the constituent a bone. Intonational prominence is assigned to the stressed syllable of the head lexical item.

5. Some problematic issues

So far, I have deliberately restricted myself to a very simple example, in order to demonstrate the feasibility of a dynamic implementation of the linking algorithms. In particular, the clause we have looked at has the simple Topic-Comment type of information structure, with the topical concept, according to Ferreira, activated first. Ferreira does not, however, make any reference to the effect of focus considerations on activation. Let us begin by looking in more detail at how RRG handles the concept of focus.

Focus types are classified, in accordance with the proposals of Lambrecht (1994), into two main types, broad and narrow (Van Valin 2005:69–70). In narrow focus only one constituent in focused, whereas broad focus includes more than one constituent. Narrow focus can be either unmarked, the unmarked position for it in English being the final position in the core of the clause, or marked, corresponding to any other position. There are also two subtypes of broad focus: predicate focus is the classical Topic-Comment type of structure, while in sentence focus the whole sentence is included in the focus domain. For the example The dog ate a bone we assumed that only a bone was in the actual focus domain, so that we have unmarked narrow focus: such a circumstance would arise if a contrast were being made between a bone and some other entity which the dog might have been thought to have eaten. But it is easy to see that there are other possibilities too: even if we maintain the dog as topical, the clause might have predicate focus, as in the answer to the question What did the dog do? rather than What did the dog eat? Furthermore, in some discourse circumstances the example might have sentence focus, with no Topic-Comment structure, as in the reply to the question What happened?

Now consider the exchange in (22), from the British National Corpus:

(22) PROF ENO: Clint! Why haven’t you changed into your kit?
CLINT: The dog ate it, sir. (BNC CHB 414–416)

Although the syntactic and semantic structures of The dog ate it is very similar to that of our previous example, differing only in the definiteness of the PATIENT, the information structures are very different. The local topic of the conversation, and so the

10. The version of the BNC used here is the original one, not the later World Edition.
most topical concept for this clause, is the kit represented by it. Ferreira's proposal would thus predict that this would be put in Subject position, giving rise to the passive \textit{It was eaten by the dog}, and this utterance would indeed have been perfectly coherent with the surrounding discourse. Instead, the speaker chooses an alternative solution, making \textit{the dog} focal by intonational means, but still opting to put it in first position in the clause. Note that although in this example \textit{the dog} represents a concept which is assumed to be available to the addressee, it is not topical: reference is being made not to some dog which has already been mentioned in the discourse, but rather to one which the addressee is expected to recognize as being, in all probability, Clint's own dog. Indeed, it would have been just as possible to use the same information structure if the speaker had wished to convey that some unidentified dog, rather than the one the addressee can presumably identify as Clint's dog, had done the eating: \textit{A dog ate it.} A corpus example is given in (23):

\begin{quote}
(23) On returning to the surface he carelessly looked away from where he had placed the brier and a \textit{snake ate it}, whole. (BNC CAC 1215)
\end{quote}

In RRG terms, what we have in each case is a marked narrow focus structure, in which focal rather than topical material is placed in initial position in the clause. Clearly, the picture with regard to the relationship between activation and positioning is much more complex than Ferreira's account would suggest.

Indeed, one recent model, the Incremental Functional Grammar proposed by Mackenzie (2000, 2004) within the overall framework of FG, claims that the first concept to be activated during production is that which the speaker wishes to focalize, and without which there would be no point in making the utterance. Material corresponding to this concept may need to be kept in a buffer while other components of the message are planned: indeed, this will often be the case, since as we have seen the unmarked position for focal material is late in the clause. In some communicative situations, however, it will make sense for the focal material to be uttered first, according to the task urgency principle of Givón (1988:275), or even unaccompanied by other, non-focal material.

It would seem, then, that there are many communicative circumstances under which it is best to place already negotiated, topical material in initial position in the clause: Halliday & Matthiessen (2004:93–94), for example, note that the unmarked situation is for what they label the Theme of a clause (the constituent which the speaker decides to take as the starting point for the message), which in English is placed in first position, to coincide with the already negotiated, given information, but that there are other circumstances in which the initial position houses new information.

A particularly attractive way of handling various types of communicative situation and their impact on the information distribution within clauses has been proposed by
Hannay (1991, 1993) within the framework of FG.¹¹ In FG topicality ("characterizing ‘the things we talk about’" (Dik 1997a:310)) and focality ("characterizing the most important or salient parts of what we say about the topical things" (ibid.)), where marked by means of some special device (word order, intonation, special constructions, particles) in the language, give rise to the assignment of the pragmatic functions Topic and Focus. Hannay suggests that the communicative effects of various arrangements of Topic and Focus correspond to different 'modes of message management', seen as planning strategies used by the speaker. These are formalized by Hannay as subtypes of the illocutionary force operator, and are expressed through the choice of (i) whether to include a Topic or not, (ii) if there is a choice of what might be Topic, then which possibility is selected, and (iii) which element is selected for placement in the initial position in the clause structure.

Although much research remains to be done before detailed proposals can be made for how such modes might be involved in language processing, it is at least possible to make some tentative suggestions. The modes presumably correspond to decisions taken by the speaker at the time of selecting concepts and organizing them into configurations which reflect the informational status which is to be accorded to each. These decisions will take into account not only what is already known by the addressee to be a current topic of the interaction, and what information can be retrieved by the hearer from sources outside the discourse itself (such as the likely referent of the dog in (22)), but also what the speaker wishes to signal as the most salient information s/he wishes to convey to the addressee. It is reasonable to suppose that this information, and not only that which is already topical, is highly active in the mind of the speaker, who must then, through the message management resources, select the configuration for the utterance which best reflects the informational statuses of the various components of the message. In terms of a dynamic implementation of RRG, modes of message management can be seen as strategies for the selection and placement of topical material and of the various types of focus proposed in the theory.

¹¹. I am indebted to Lachlan Mackenzie for pointing out to me the usefulness of Hannay's proposals in the context addressed in this paper. Further discussion of Hannay's proposals can be found in Butler (2003b:87–90, 96–97).
6. A more complex example of dynamic implementation of the linking algorithms

In the light of the foregoing discussion, let us now look at how the dynamic implementation of the RRG linking algorithms might work with the more complex example given in (24):

(24) What did you give Mabel for a birthday present? (BNC FA5 1134)

The context of this utterance in the corpus from which it is taken suggests that Mabel is known to both speaker and addressee, and the speaker knows that the addressee has given Mabel a present, but not what that present was. As a conceptual representation I suggest something along the following lines:

(25) \textit{ASK (TRANSFER/UNMARKED; (ACTIVE/SINGLE/DEFINITE/ANIMATE/AFFECTER/TOPICAL/ADDRESSEE; BRAND-NEW UNANCHORED/INDEFINITE/INANIMATE/AFFECTED/FOCUSED; QUESTIONED ENTITY; ACCESSIBLE OR INACTIVE/SINGLE/DEFINITE/ANIMATE/RECIPIENT: MABEL; ACCESSIBLE OR INACTIVE/SINGLE/INDEFINITE/INANIMATE/PURPOSE: GIFT/BIRTHDAY; PAST))}

Here, the specification \textit{TRANSFER/UNMARKED} is intended to convey that the speaker conceptualizes the process of transfer more narrowly as an unmarked one of giving, rather than presenting, donating, lending, etc., all of which would be differentiated by the addition of further qualificational concepts to the basic concept \textit{transfer}.

On Ferreira’s assumption that the already most highly activated concepts are the first to be processed, we might expect the process of concept activation to begin with the \textit{addressee} concept, which is active throughout an interaction. This activates the template for a NP with pronominal head, as follows:

(26) \text{NP} \\
\text{PRO} \\
\text{you}

However, we have seen that it is reasonable to suppose that concepts which are under focus will also be strongly activated in the mind of the speaker. In the present case, this

\begin{itemize}
\item[12.] The semantics-to-syntax and syntax-to-semantics mappings for this example are discussed in some detail in Butler (2003a:144–149), though without any attempt at dynamic implementation.
\item[13.] Note that NPs headed by pronouns or proper nouns do not normally have a layered structure.
\end{itemize}
means that the questioned entity concept, which represents the information which is being sought, will also be highly active. This concept, in combination with the illocutionary force operator ‘interrogative’ which arises in response to ask in the conceptual representation, immediately calls up a template for the clause with a pre-core slot, since in English a wh-item goes into this position. Furthermore, the properties of the questioned entity (inanimate, affected) ensure that the lemma activated to fill the pre-core slot is what, rather than for example who, why or where, as shown in (27).

(27) CLAUSE
    PrCS  
    CORE
    NP
    PROWH
    what

Furthermore, since the concept for the process is not one which can be realized by be, the rules of English specify that a form of do must be inserted between the wh-word and the first core argument. We also know that the value of the TENSE operator must be ‘past’, giving did as the form to be inserted.

The material so far activated is placed in a buffer while the concept transfer/unmarked corresponding to the predicate activates the lemma for give, whose logical structure is then retrieved from the lexicon:

(28) \([\text{do}^\prime (x, \emptyset)] \text{CAUSE} [\text{BECOME have}^\prime (y, z)]\]

A great deal of information immediately becomes available as a result of this retrieval, due to the operation of the semantics-to-syntax linking algorithms. The Actor-Undergoer Hierarchy specifies that the Actor will be the first argument (EFFECTOR) of do’ (i.e. x). It also specifies that the default for Undergoer will be the second argument (PATIENT) of the stative have’ predicate, since of the available arguments this is the furthest right in the hierarchy – this assignment is seen in the clauses in (29) and (30):

(29) Jim gave a CD to Mabel for a birthday present.
(30) What did you give to Mabel for a birthday present?
However, with certain verbs, English allows an alternative assignment to Undergoer, the argument next left in the hierarchy, i.e. the first (y) argument of have’, resulting in clauses such as:

(31) Jim gave Mabel a CD for a birthday present.

and our target clause as in (24). In order to find out what might motivate the choice of one structure rather than the other, careful corpus-based studies such as that reported by Siewierska & Hollmann (this volume) are needed.\footnote{Siewierska & Hollmann (this volume) show that in the Lancashire dialect from which data are taken for their study, a third possibility, with the Theme constituent as the first object in the double object construction, is attested, and that when both objects are pronominal this order is actually more frequent than the variant with the Recipient first. I shall not discuss this possibility further here.} In our present case, the speaker, for whatever reason, opts for the second possibility. So we have: x = Actor, y = Undergoer. We also know that the x argument of do’ is an EFFECTOR, the y argument of have’ is POSSESSOR and the z argument POSSESSED.

As we have an active clause, and English is an accusative rather than an ergative language, we get the default assignment of Privileged Syntactic Argument to the Actor, so: x = Actor = PSA. Further, the rules of English specify that the PSA is linked to the first core element of an active clause.

The formulator now needs to work out the appropriate structure to be inserted into the CORE slot of the PrCS template. There are three distinct specified argument positions in the LS, so there would normally be three arguments in the syntactic structure, according to the syntactic template selection principle given in Section 2.

But English has a language-specific rule saying that if there is an argument in the pre-core slot (PrCS) the number of arguments in the core is reduced by one (Van Valin 2005:130). We therefore need a core template with 2 argument positions, i.e. the same Core-3 template given as (15) earlier. This core template fits into the CORE slot of the clause template, and the latter into the sentence template (there is nothing in the LDP in our example). So far, then, we have the structure in (32).
We may hypothesize, following Ferreira, that the material planned so far may be sent to the phonological encoder at this point. Note that in different communicative circumstances, the generation of the clause could have stopped here: if it were clear to speaker and addressee that the conversation was about the addressee giving something to someone, a perfectly proper formulation would be *What did you give?* In our case, however, these circumstances do not obtain. The concept *Mabel*, either immediately accessible or currently inactive and therefore in need of reactivation, invokes the template for a NP with the proper noun *Mabel* as head:

(33) \[ NP \]

\[ N_{\text{prop}} \]

\[ Mabel \]
We now have a complete linking of arguments to the syntactic structure, as follows:

\[(34)\]

\[
\begin{array}{c}
\text{SENTENCE} \\
\text{CLAUSE} \\
\text{PrCS} \\
\text{NP} \\
\text{PRED} \quad \text{V} \\
\text{PRO}_{WH} \quad \text{PRO} \\
\text{What} \quad \text{did} \quad \text{you} \quad \text{give} \quad \text{Mabel}
\end{array}
\]

Once again this might, in other communicative circumstances, be enough, but our conceptual specification also includes the purpose of the transferred entity as gift/birthday. This latter concept triggers the structure for an NP with the lemma *birthday present* (treated here, probably unjustifiably, as a single compound item) as head:
The **purpose** concept in the specification of the referent for this NP triggers the marking of this NP with the preposition *for* (the conceptual specification role would have triggered *as*), so that the following PP is constructed:

(36)\[ PP \]

```
P
  NP
    CORE_N
    NUC_N
    N

  for

  a

  birthday present
```

INDEF \rightarrow NP
This can now be slotted into the free PP/ADV slot for the peripheral adjunct, giving the final structure as follows:

(37) SENTENCE
    | CLAUSE
    PrCS
    NP NP NUC NUC
    PROWH PRO PRED V Nprop P NP NP
    What did you give Mabel for a birthday present

Actor = EFFECTOR = PSA
Undergoer = POSSESSOR

for’ (w, [do’ (x, ∅)] CAUSE [BECOME have’ (y, z)]
7. A note on the possibility of parallel processing

I have not pursued here the possibility that in addition to the linearisation process which is clearly required in language production, there may be some involvement of parallel processing mechanisms (see Fortescue, this volume). We have seen, however, that RRG postulates the involvement, in the mapping between semantics and syntax, of constructional templates which contain morphological, lexical, syntactic, semantic and pragmatic information, and that in the process of constructing linguistic output these templates are fitted together in appropriate ways. This process is effectively one of constraint satisfaction, and the theoretical mechanisms of RRG therefore offer the possibility for parallel generation of certain parts of the utterance and the combination of these parts in such a way as to satisfy all current constraints or to achieve an acceptable compromise among them. This possibility remains a topic for future investigation.

8. Conclusion

The principal aim of this paper has been to show how the mapping between semantics and syntax proposed in Role and Reference Grammar can be implemented in a dynamic manner which takes account of evidence from work in psycholinguistics. This work is intended as a contribution to a much larger and more ambitious research project sketched out in Butler (2003a:489–499), for the construction of a model of language and language use, with a dynamic structural-functional grammar at its heart, responding to a whole set of criteria of adequacy, not only psychological/cognitive, but also discoursal, sociocultural and acquisitional, as well as to typological concerns. As discussed in Butler (2003a, 2003b) RRG has already shown itself to have high levels of typological and acquisitional adequacy, but has so far paid very little attention to sociocultural matters and has restricted its treatment of discourse to those aspects of discourse pragmatics which are closely bound up with the grammar. It remains to be seen whether the proposals made in the present paper will sit well within a larger model which is sensitive to these other important concerns.

References

Fortescue, M. (this volume). The non-linearity of speech production.
Christopher S. Butler


Incremental implementation of RRG algorithm


Grammar, flow and procedural knowledge
Structure and function at the interface between grammar and discourse

Peter Harder
University of Copenhagen

This contribution addresses the interface between grammar and discourse. It approaches grammar as a socially recognized procedure for coding, with a cognitive basis in individual language users. However, it emphasizes the need to distinguish the cognitive representation of grammar from the flow of discourse itself, which involves a more complex socio-cognitive process. Within this framework, grammatical structure is presented as a result of functional differentiation in communication, and grammatical slots emerge as basic functional-structural interface units between grammar and discourse. The layered structure of the clause that has been posited by Functional Grammar is then seen as forming the basis of this view of grammar as procedural knowledge, and emphasis is placed on the inevitable simultaneous functioning of all layers in grammar. The article finally turns to the relation between this view of grammar and the role of incrementality in the flow of discourse, with specific attention being paid to formulaic language and units of discourse.

1. Introduction

The relation between grammar and discourse is a basic question for all functional approaches to grammar. In this article, I am going to take up this issue in relation to the discussion that has taken place in the context of the development from Functional Grammar (cf. Dik 1989, 1997) into Functional Discourse Grammar (FDG), cf. Mackenzie and Gómez-González (Eds., 2004). The feature that I am going to focus on is the close association in the model between the descriptive system and the online production process in the mind of the speaker. The issue involves a number of foundational questions, including the question of ‘psychological reality’ and the relation between ‘pattern’ and ‘process’ (cf. Fortescue 2001, 2004). In addition, there is also a question of clarity: there has been some disagreement on exactly what status FDG has in relation to that dilemma. My discussion, however, does not depend on the details of the model; central to it are really only the following three features:
The first is that FDG operates with three major 'levels': the interpersonal, the representational and the 'expression' (or 'structural') level. All three are conceived as describing a 'linguistic unit', with a directionality of description (inspired by Levelt 1989) such that the interpersonal level constitutes the point of departure and captures the 'act' properties (assertion, summons, etc.). The interpersonal level may then 'draw upon' the representational level (cf. Hengeveld 2004a:11) with the proposition ('communicated content') as the highest unit – and both interpersonal and representational choices are then realized by the morphosyntactic structure of the clause, at the 'expression' level.

The second is that such a 'top-down' descriptive order (from the whole act down to morphosyntactic details) is not supposed to mirror the discourse process, or the psychological processing, point by point (cf. Mackenzie 2004; Hengeveld 2004b). The question I am addressing is: If we want to have a theory that captures the basic relationship between structure and flow, without naively assuming that they are one and the same thing, how exactly should they be linked?

The third point is that in FDG, the inspiration from Levelt's speech production model is reflected in the fact that several authors build some form of top-down temporal flow into their understanding of the model (cf. Hengeveld 2004b; Mackenzie 2004; Bakker & Siewierska 2004). An intuitively attractive version is the 'tiles on a roof' analogy, in terms of which lower levels are temporally delayed, but overlap with the higher levels: The interpersonal level goes first, then the representational layer begins, and after a slight delay expression rules start to operate, such that the three levels interact for a stretch of time (cf. Hengeveld 2004b:368). I believe this seductive analogy is wrong, and that the fault in it has to do with a basic misunderstanding of the relationship between flow and structure. In this article I argue for a solution that has some of the 'intermediate' properties that the model aims at (Hengeveld views it as a pattern model, but one that reflects the process, cf. Hengeveld 2004b:366) – while maintaining a total separation between grammar and actual flow. In advocating such a radical separation, I differ not only from the positions of Hengeveld, Mackenzie, Bakker and Siewierska as cited above, but also from Butler (this volume). The reason for this is not desire to maintain a tidy theoretical distinction, but because I think the real world is constituted in such a way that the two are different types of thing. The following properties are central to that picture: (1) grammar is a basically a socially recognized procedure for de- and encoding messages (cf. Harder 2003), reflected in the speaker's mind as 'procedural knowledge'. (2) Procedural knowledge is different from actual processes in that it is not part of the discourse flow, and different from the standard representational view of grammar in that it does not specify a product, but a process. (3) Grammar arose evolutionarily as a result of a 'big bang' that splintered the holophrase, i.e. created an extra level of complexity in relation to previous forms of communication where each (unstructured) signal encoded a complete (unstructured) 'utterance'. The home ground of grammar is thus utterance-internal structure, and grammatical relations between utterances must be understood on the basis of utterance-internal grammatical properties, especially content-based de-
pendency relations. (4) In spite of the progress made by FDG (cf. Harder 2004), I think it is desirable that the further development of the model should aim to facilitate the explicit representations of links between the three levels (interaction, representation and expression), instead of allowing the three strata to drift apart, suggesting that some are more flow-like than others. I discuss especially subject assignment and P1-filling as choices where expression and content should be more clearly linked, also as prime examples of how grammatical options interface with discourse factors (rather than iconically reflect them).

The notion of layered structure has been at the core of this issue. In FG, relations with discourse used to focus upon ‘higher’ and ‘bigger’ layers taking over outside the sentence boundary. Among the problems with this is that the lowest level of discourse layering does not coincide with the highest layer of clausal layering, cf. Kroon (1997). In addition, an elaborate (or, as Lachlan Mackenzie once expressed it, ‘generously upholstered’) layered grammatical structure, with discourse adding even more elaborate layering, is potentially misleading with respect to the typical coding complexity of discourse instalments. In focusing on the basic status of holophrases, as well as on the importance of the incremental perspective in understanding the progress of discourse also within domains that lend themselves to purely grammatical description, Lachlan Mackenzie (1998, 2000, 2004) has made a major contribution to putting research into the relation between grammar and discourse on the right track.

The points I announced above, however, also have some bearing on the incremental dimension that Mackenzie has rightly posited as central to this discussion. While Mackenzie has used the adjective about the grammar (‘Incremental Functional Grammar’, cf. Mackenzie 2000), I argue that this property is essentially a discourse property while grammar operates to suspend it – even though discourse pressure makes itself felt also within the clause. Quoting Jackendoff (1997:7–8), Mackenzie argues that FDG combines his two last positions: (b) one can maintain that processing mechanisms can ‘consult’ or ‘invoke’ a declarative grammar; (c) or one can claim that the processor embodies the grammar, i.e. that grammar is itself procedural. While I agree with the word ‘procedural’ as a description of the status of grammar, I understand it in such a way that grammar does not model actual production. With respect to the point where Mackenzie (2004) discusses grammar as limiting incrementality, I argue that Mackenzie’s account, though true as far as it goes, does not sufficiently take account of one dimension of the issue. He sees the role of grammar, in so far as it is non-incremental, mainly as an impediment to the free incremental flow of discourse, instead of highlighting the functional potential of the fact that grammar rises above ‘hand-to-mouth’ encoding. Since this is bound up with the basic nature of grammar as I see it, I am going to begin the argument by outlining in some detail the approach to grammar I adopt.
2. Linguistic structure as a result of functional differentiation

Evolutionary ‘scenarios’ have become increasingly popular in recent years. One motivation is that a ‘before’ vs. ‘after’ picture highlights those properties that make a significant difference more strikingly than a description in terms of the way things are after the change has occurred. The version I defend owes some features to Deacon (1997, 2003a, b). In terms of evolutionary development, structured language can be understood in analogy with other structured phenomena in the biological world, like courtship displays or webbed feet – a new thing that came into the world at a certain point in an evolutionary course of events. Human languages differ from other systems of animal communication in multiple ways, but one of the two main points of departure for understanding the nature of the difference lies in seeing the pre-human stage as characterized by a unity between the analysis into signals and acts. Each signal in the animal’s repertoire conveys a complete act. Alarm calls, barks of submission, signals of aggression or fear – all have an expression side which is unanalysable into component expressions as well as a content side that is similarly unanalysable, and which slots directly into whatever course of events the signals form part of.

The interest in outlining such a semi-fictive point of origin lies in seeing precisely what happens when a human language evolves in a context of that kind. In terms of the structure of reality, the most central event, I suggest, is the rise of a new level of analysis in sign systems (as part of evolving reality): in addition to seeing each sign in the context of the situation to which it contributes, we now need to see each sign in the context of the other signs that are part of the linguistic act to which the sign contributes. To take an obvious example, words like the or of cannot be understood directly in relation to the situation in which they are used – they need to be understood as serving a function in relation to those signs with which they collaborate in coding a whole linguistic act. Such collaboration is standardly understood in terms of composition. What I am concerned to stress here is that in terms of evolutionary directionality, composition presupposes a logically prior step of differentiation. Before the uniformly holophrastic pattern started to crumble and utterances began to be divisible into sub-utterance fragments, the question of composition did not arise. But it does arise the moment differentiation becomes possible, as part of the very same event: fragments that cannot be combined into whole utterance meanings are no use to the speaker. So flexibility has a price: human languages come with the label ‘some assembly required’.

Something is lost here, also communicatively. It is simpler when there is one specific call for each type of relevant occasion, and hence no problem in linking calls with the situational events they interact with. In that sense, ‘functional innocence’ is lost when grammar arises. But the functional potential of language obviously becomes vastly greater with the loss of innocence and simplicity. As pointed out by Deacon (1997), this is especially due to one thing that happens on the developmental path from communication by means of calls to communication in a human language. With ‘symbolic’ meaning, Deacon highlights the rise of meaning that is decontextualized, i.e. anchored not in the immediate situation but in a universe of generalized meanings.
abstracted from both iconic and indexical links with the environment, reflecting the design feature of ‘displacement’ (cf. Hockett 1958:354 and also Bloomfield 1933:30). In the terminology I use, the word ‘conceptual’ is reserved for what Deacon calls ‘symbolic meaning’: a ‘concept’, in the relevant sense of the word, stands above any concrete instantiation, subsuming actual as well as potential specimens. The development whereby human language became able to tap a universe of ‘eternal’ conceptual meanings is generally regarded as the main difference between human and animal systems of communication.

There is thus a generally recognized link between representational, conceptual ability and grammatical structure. In FDG, this is reflected in the fact that grammatical structure is mainly associated with the representational level, while unstructured signals go directly from the ‘interactive component’ to the ‘expression’ component. Within generative grammar, Newmeyer suggests that language structure is based on (pre-existing) conceptual structure (cf. Newmeyer 2003), rather than on pre-existing forms of communication, so that syntax is grounded in the structure of representational concepts. Such a direct link between representation and structure, however, is at risk of underestimating the functional role of clause structure. It is essential to the position I defend that the functional-interactive element in clause meaning is part of the structural differentiation into less-than-full-utterance meanings. A tendency to put interaction ‘above’ structure would therefore render FDG a less ‘functional’ model of grammar.

3. Slots as basic functional-structural interface units: External and internal function

In order to understand the combined role of structure and conceptualization in enhancing the functional potential of human language, it is illuminating to look back at Chomsky’s seminal critique of Skinner’s behaviorist understanding of language (cf. Chomsky 1959). The behaviourist world picture allowed only simple cause-effect relations, corresponding to each signal having its own effect, with no room for internal structure intervening between signal and response. In a word-based account, that was what gave rise to the Markov-chain (finite state) model: each word had an effect which was allowed to depend on the previous word but could not reach beyond that. Chomsky (1957) argued that understanding and encoding linguistic utterances could not be accounted for on a word-for-word basis, and I think that point still holds good. Chomsky thought what was missing in the behaviourist view was hierarchical, formal structure. However, the central arguments for structure-dependence have always involved the content side as well: what verb do you front in order to encode an interrogative version of the declarative clause; and how can you tell who does what to whom? Structure-dependence, from a functional-grammatic viewpoint, involves the way word meanings collaborate in bringing about a total meaning that is more than the simple sum of its parts. Because the meanings of linguistic items are defined in
terms of the way they contribute to a whole clause meaning, there is a close correlation between what the items mean and what jobs they have in the whole clause meaning. Verbs predicate activities, abstracted away from participants – and therefore they occur in the company of expressions denoting those participants that the verb itself does not designate (and standing alone, they foreshadow what participants they can go with). Modal adverbials like perhaps designate epistemic status, and need to take a proposition in its scope, to which that epistemic status can be assigned. This type of semantic dependency relation (with Langacker’s term) is crucial both in understanding the ‘fragment’ status of differentiated item meanings, and in understanding what happens when we understand them on their own.

But not everything can be projected from item meanings; they can also change to fit into a larger whole in new ways. Dual word class membership in verbs and nouns is pervasive in English, rendering e.g. time flies ambiguous (cf. the reply ‘you can’t – they fly too quickly’). In such cases, we need to recognize both the item and the particular ‘job’ or ‘slot’ into which it fits in order to understand the whole meaning, and the contribution of the parts. Below, I shall use the slot/filler dichotomy about this duality. Others, including the tagmemic theory, have viewed this relationship as pivotal in understanding language. The ‘filler’ properties are what items bring to the utterance in which they are included. Tagmemics was a source of inspiration to Dik, cf. Anstey (2004:29), and the sense of ‘function’ that motivated Dik’s christening of his theory, is bound up with the three sets of functions (semantic, syntactic and pragmatic) which serve as slots for NPs (‘terms’). The slot properties are not due to the filler in itself, but are associated with the position in which it is used, its ‘role in the larger whole’ that it is used to fill.

Harking back to the basic discussion of ‘function’ as a relationship between a behaviour and the environment, we can now see that the rise of internal slots in an utterance necessitates a two-level approach to ‘function’. In terms of the direct relation with the environment, the operative level of analysis is the whole utterance (half an utterance is not a relevant unit of analysis). This follows from the fact discussed above that the utterance is the unit of speech. When there is a situation that calls for communicative action, what we aim for is to make the right utterance to serve the need. If the utterance has the function we aim at, we are communicatively successful.

But given that we have the option of using a level of coding choices below the level of the utterance, the functional question arises also at that level. There is a job we need a word to do, and if we choose the right word, we solve the problem (provided the rest of the utterance does its job). But the function of the word is defined in terms of its job in relation to the rest of the utterance. The word cold does not have a uniform utterance function. It helps to achieve different situational purposes for us, depending on whether we insert it in the slot it is ____ in here and please get me a _____ beer. The choice of the word must be understood ultimately as motivated by the actual discourse purpose of the speaker – but she cannot use the word cold to achieve her discourse purposes simply by virtue of knowing about the relation between the expression cold and the situational context (as she could in the case of, e.g., the expression hurrah!). In the
one case, cold helps to convey a complaint about room temperature, and in the other it helps to specify a request. The point is that the functions of sub-utterance linguistic elements depend saliently on the overall (content-syntactic) structure constituted by the hierarchy of linguistically coded meanings, rather than only on the situation and the item meaning in itself: structure mediates between individual items and overall utterance function.

This understanding of ‘structure-dependence’ places it as a direct consequence of the differentiation of the whole utterance into independently codable slots with functional relations between them. An utterance as a whole also fits into a slot, but that is defined in relation to the external features of the situation; in conversation, a ‘turn’ would be a good approximation to an utterance-level (or move-level, cf. Hengeveld 2004a), externally defined ‘slot’. But inside the macro-question of how to select utterances that fit into the external situation, we have the micro-question of how to select linguistic elements that together constitute an utterance with the right semantic properties. We may choose to collapse the two questions into one by selecting an utterance-level signal as a whole. But the procedure of encoding messages in a human language is constituted so as to raise both questions.

The reason why this can be understood in terms that are fully functional, even while it stresses the importance of distinguishing between utterance-external and utterance-internal function, can be explained by an analogy with another type of functional whole, namely the business company. The central points include the existence of a two-way developmental process, where the basic direction is that of differentiation, but the success of this operation depends on the skill of integrating the differentiated elements into larger wholes. The developmental and dynamic perspective in both cases perseveres as a feature of the ‘synchronic’, static state of the complex phenomenon under investigation.

The starting point for the process is the event of one man (let’s call him Jones) setting up a business for himself. As long as the man works his company single-handedly, all relations with the world meet in one and the same unit, namely himself. The functional relations with the world may be multifarious and complex, as they are in the case of communication, but they are not matched by a corresponding structural complexity. Jones is the company; the company is Jones. An external-functional description, i.e. a description of what the company does in relation to its environment, will be an exhaustive description of Jones Inc.

Let us now suppose that the company is a success. Jones gets paying customers to such an extent that it is no longer possible for him to do all the work on his own. He therefore gets someone to join him in the company. This is the situation that I want to present as analogous to the rise of complex utterances. We now have two issues instead of one. Previously the questions: ‘what should the company do?’ and ‘what should Jones do?’ boiled down to one and the same question. Now the two questions are no longer the same thing, because in addition to deciding what the company should do, it is necessary to decide what Jones should do and what the new man (Smith) should do. The internal division of labour cannot be derived straightforwardly from the external
function of the business company. Asking about what the job of the new man is going
to be raises two sub-issues that are linked but irreducibly different (i.e., one is partially
autonomous of the other): How do Smith’s talents fit with what the company as a
whole does, and how do they fit with what Jones does?

Up to a certain point of complexity, the internal organization of a business com-
pany may maintain functional ‘credibility’, because the internal organization is ‘trans-
parent’ in relation to the external functional demands upon the utterance. In terms of
‘iconicity’, the internal structure may come across as but a shadow of the functional
organization of the communicated message. If a two-man company is organized so
that Jones does production, while Smith does sales, internal structure reflects exter-
nal function. Similarly, when a sentence has two structurally signalled constituents,
subject/topic/referent and predicate/comment/ascribed property, and these are in fact
the topic and comment of what is being discussed, clause structure is functionally
transparent.

As the organization becomes more complex, however, links between internal and
external function will not continue to be transparent and obvious. There will be a need
of ‘managerial’ functions in order to keep the organization coherent. The people who
are doing production in large companies typically look askance at the proliferation of
management, but if co-ordination between sale and production fails, it will be bad for
the whole, even if both on their own are more transparently motivated than manage-
ment. How much management is needed is an open question (which is to some degree
analogous to the question of the functional utility of elaborate case and agreement
marking). In any case, the degree of motivation cannot be established unless the con-
stituent is viewed in all its relations with the whole meaning to which it contributes.
The ‘some assembly required’ condition that goes with human language involves a
complex network of links between items, mediated by the slots that these items can fill.
And only via the internal relations can we understand the external functional relations
that the complex utterance as a whole can establish.

4. Layered clause structure as procedural knowledge

The relationships between content slots (‘jobs’) are to a considerable extent describable
in terms of the layered structure of the clause (the major exception being information
structure, about which more below). The basic logic of layered structure reflects a type
of ‘process’ understanding, but one that is different from online, real-time process-
ing, because it is defined in terms of a task that has to be carried out. In that sense,
cf. Fortescue (2004, Note 1), it places itself apart from the basic distinction between
pattern and process that was discussed above. I have earlier described it in terms of
‘recipe’ (as opposed to ‘cooking’) or ‘compilation’ (as opposed to execution), cf. Harder
(1996:214); the same idea is what is inherent in viewing a task from a ‘procedural’
point of view. This description has acquired some renewed relevance because of the
ambition of FDG to be a ‘pattern-that-reflects-process’ type of description.
Rather than begin inside the mind of an individual in the Chomskyan fashion, I think it is more illuminating (cf. Harder 2003) to take the point of departure in language as a collective entity. Language as the property of a speech community constitutes what with Gibson’s term (Gibson 1979) can be described as an ‘affordance’ for speakers in that community. The options offered by the community can be regarded as a set of ends-means relationships, as available procedures for exchanging complex messages. As children grow up and acquire the language, they ‘adapt’ to the linguistic environment by acquiring a cognitive apparatus (a ‘procedural’ knowledge of the language) that enables them to make use of the linguistic affordances of the community. In this context, I would like to put forward this view as an attractive interpretation for the new FDG model.

This differs from the way grammar is standardly understood, being thought of as ‘declarative’ in contrast to the discourse-oriented ‘procedural’ elements (cf. e.g. Mackenzie 2004, who specifically reserves the word ‘declarative’ for representational clausal structure as opposed to the interactive and expression levels). In the computational source of the procedural/declarative dichotomy, however, the basic idea is the difference between a program specifying what to do in order to produce a representation and the finished representation as such. It is not really a question of the inherent nature of the object – since the procedure (in a well-made program) uniquely specifies the result, the two notations are effectively equivalent. If we can state a computationally precise ‘declarative’ structure for a clause, we can also choose to represent it in procedural form, with no empirical difference. Thus procedural capacity is not bound to represent only tacit knowledge of the ‘knowing-how’ type, but can also be applied to skills that are accessible to full conscious awareness (cf. Johnson-Laird 1983:248).

The point of thinking of clause structure in terms of procedural knowledge is that a procedure has the interesting feature of being neither an event in time nor a wholly inert construct. Procedures can be held in static readiness, but they can also be invoked once we are ready to move. Procedural capacity is naturally thought of as being implied in hosts of other skills than language, such as in knowing how to construct complex artefacts and in knowing how to conduct a social event such as a trial. To return to the business analogy, the most functionally central property of a production company is its overall capacity (irrespective of how it is cognitively distributed) to take the relevant input and use it to produce the desired output. A declarative specification of the output is less obviously useful than a procedure for making it, if the procedure guarantees the right product. The ‘conveyor belt’ metaphor used of ‘old FG’ fits nicely in here, once it is understood that the relevant properties are those that the facility has also when the belt is not actually up and running.

Procedures involving layering, moreover, have another interesting property. They are set up in order to ensure a desired goal (e.g. a car) – but at the same time the inherent logic of the procedure does not begin with the car, but with the necessary initial input that is then subject to a number of operations before the car emerges as the end goal of the procedure. Nor is this an exception due to the mundane nature of car-making. Even if the aim is to build a cathedral with a spire on top that seems to
touch the heavens, the actual production process has to begin with digging a hole in
the ground, proceeding in a 'bottom-up' manner from there. Applied to the case of
utterances, this raises a question of whether a basically top-down procedure, such as
the one embodied in FDG, needs to actually work strictly top-down. As also discussed
in Mackenzie (2004), the “tiles on a roof” procedure entails that one stage can trigger
the next before being completed – which allows for a partial bottom-up element when
ends have to meet. Speech errors demonstrate that ends may in fact not quite meet
when processes at different levels are running simultaneously, as in *she come backs
tomorrow, the quake caused extensive valley in the damage* (cf. Aitchison 1998:257),
where the words have been recruited and also the syntactic structure, but the words go
into the wrong slots, short-circuiting strict top-down planning. As an example of the
same thing, Bakker and Siewierska (2004) also mention that phonological elements
may be recruited at a fairly early point in the overall procedure. It seems plausible to
assume that the overlapping of various stages should be interpreted so liberally as to
allow a certain measure of bidirectionality, i.e. one whereby the initial awareness of
the goal may prompt a process of gathering the necessary raw materials, which are
subsequently processed in order to produce the relevant output.

Tommasello’s theory of language learning (cf. Tommasello 1992, 2003) suggests that
‘verb islands’ play a significant role, i.e. that verbs are learnt with a particular ‘predicate
frame’ and persist in that form for a long time and are only slowly subsumed under
more abstract patterns. Also the learning theory of Pienemann (cf. Pienemann 1998:7)
suggests that the language learner needs to approach clause structure bottom-up: being
able to handle lower levels is a precondition for being able to handle higher, more
complex levels. Although learning and production are different processes, the existence
of a bottom-up path in learning suggests an ease of access that might be expected to be
useful also for production purposes.

Procedurally, layering is a matter of input-output relations. The logic of the ‘su-
perimposition’ process begins with arguments, then applies the predicate to it, then
adds the propositional operators and satellites and finally cap it with an illocution
and integrate it in whatever higher-order discourse context we recruit it for. If we are
speaking in terms of procedural knowledge as opposed to an actual execution of the
procedure, there is no need to choose between a top-down path and a bottom-up path:
knowledge may be organized so as to make both available. The structure of the state-
ment *Germany lost the match yesterday* can without contradiction be compiled in both
directions: the rough top-down path is: assertion (ill), about the past (tense), more
specifically ‘yesterday’ (temporal satellite), with an ascription of the predicate *lose*
to the arguments *Germany* and *the match*. Bottom-up is simply the other way: first we
create a basic state-of-affairs by predicating *lose* about the arguments *Germany* and *the
match*, then we place the losing event yesterday, assign past tense to it and use it to
express a statement. As long as it is understood as knowledge rather than actual pro-
cess, the essential thing is the inherent relationship. A sufficiently advanced assembly
line can in principle be furnished with buttons that can be pressed at any point in the
procedure – the procedural knowledge itself makes the buttons available, but does not press them. Only a concrete discourse event does that.

5. The limits of incrementality

Incrementality is pervasive in what Saussure calls langage, i.e. the phenomenon of language viewed in abstraction from the distinction between system and use. In speech production this is so for at least two reasons: it unfolds in time where one thing has to come after another – and taking one step at a time is the simplest procedure there is. Mackenzie (1998, 2000, 2004) shows how many characteristics of natural speech can be captured by reference to this basic pattern. A central fact is the frequency of holophrastic and set phrases as against productively combined complex clauses (see also Section 7). In addition to this, however, Mackenzie (2004) shows that even within complex clause structure, there is a gravitational force towards taking words in a step-by-step fashion. He shows also that there are limitations, which he discusses mainly in terms of ‘constraints’ that are built into what he calls the ‘declarative’ grammar that is associated with the representational level – such as the need to express the patient in relation to the verb eat (otherwise the verb would mean ‘have a meal’), even if the focus is exclusively on the verb ‘eat’ itself. Also, in languages like English, there are fairly strict rules for word order that limit free incremental expression of words – whereas in other languages,

...the temporal order in which the concepts become available may be much more directly reflected in word order, with the declarative grammar serving to ensure above all that the semantic relations between the constituents remain transparent

(Mackenzie 2004: 189)

In the following, however, I would like to argue that the idea of concepts becoming available in a particular order should not be understood so as to make each conceptual instalment “incremental” in the same sense that actual discourse acts are incremental.

The reason for this has to do with the second point in the quotation above, namely the role of the grammar in keeping semantic relations transparent. In terms of the metaphor of division of labour between fragments of clause meaning that I introduced above, there is a hierarchical logic built into semantic relations between clause members which means that grammar operates at a level where all concepts are in some sense “available” at the same time. The interesting thing, which is also central to determining the precise mode of existence of grammar in a functional linguistic theory, is to determine precisely in what sense all concepts have to be simultaneously available.

An example of this is the relation between premodifying adjectives and nominal heads, as in the phrase

(1) pregnant females
An observation by Dahl (1971), quoted by Dik (1989:116), shows that the information in the head has some form of “logical priority” in relation to the modifying adjective. If not, there would be no difference between paraphrasing (1) as ‘individuals who are female and pregnant’ and ‘individuals who are pregnant and female’. But the latter is odd, because once you have ‘pregnant’, it is redundant to add ‘female’. This observation is relevant for linearity in that (1), from a purely linear point of view, exemplifies the odd and redundant order. Nevertheless, there is nothing wrong with it – because the information in the head word females is already mentally presupposed when the word pregnant is used. To spell it out in relation to the quotation above, in expressing the concept ‘pregnant’, the speaker is not simply incrementally adducing one relevant concept, subsequently adding ‘females’ when that idea in turn becomes available to him.

I am not suggesting that Mackenzie disagrees with this – but it makes explicit a limitation in the linear point of view, one that is not expressed solely by pointing out ways in which ‘declarative’ grammar prevents ideas from coming out in a psychologically natural order. This might be argued to involve an arbitrary constraint on word order of the kind that Mackenzie focuses on, arbitrarily putting adjectival modifiers first in English. On closer inspection, however, that consideration cannot take care of the whole issue. Although in English you cannot reverse the order (because there are restrictions on putting adjectives in head position), in Danish you could use the equivalent of ‘female pregnant’, [kvindelige gravide], making the female idea come out first and the pregnancy afterwards – but it still sounds weird because the hierarchical logic overrides linear ‘naturalness’. The point in this exercise is that the grammar of Danish allows whatever linear order may be psychologically natural, but nevertheless, within the domain of grammar, linear relations are secondary to the relations that meanings have in relation to logical hierarchy of the whole clausal meaning. Speakers whose cognitive apparatus only allows them to handle half a message content at a time would be communicatively crippled.

This does not argue against Mackenzie’s observation that in some languages it is possible to express ideas in a pragmatically and psychologically natural sequence (as in the Prague School directionality from less to more communicative dynamism). However, in the present context the point is that linear order is secondary to the hierarchical order reflecting the division of labour between the semantic constituents in the clause (as signalled, in the Slavic languages, by affixes rather than word order). Unless the speaker was capable of relating all the elements to a simultaneously available whole meaning, he would not be able to understand what the message was. So that kind of ‘free’ word order does not reflect a basic ‘step-by-step’ progress, but a linearization of a presupposed, complex whole.

If the actual production process is the presupposed object of the description, the notion of simultaneity that I have invoked is problematic. It is perfectly possible to begin a sentence with no clear idea of what the whole message is going to be. Academic readers of David Lodge’s Small World will remember and commiserate with the lecturer who could not get beyond the adverbial Clearly on the top of the second page.
in his manuscript ("yes – but clearly what?"). But in that context ‘incremental’ is a property of the (disastrous) production process. The grammatical relation between the epistemic adverbial clearly and the following proposition, just like the grammatical relation between premodifier and head, presupposes a universe where both elements are available at the same time, and if the presupposed element is missing, the presupposition failure will have consequences at the discourse level. Seeing grammar as constituted by socially available procedures, internalized as procedural knowledge, provides a plausible format for describing this non-temporal relationship. The English language offers the possibility of designating a complex category by selecting a general category as the nominal head and using a premodifier to delimit the relevant conceptual subdomain, and also to signal the epistemic status of a proposition by prefixing an adverbial disjunct. Whether the speaker rashly starts off with Clearly... without having a suitable follow-up in mind is irrelevant for the description of the grammatical relationship. So when Mackenzie suggests that "incremental processing in the production of utterances works on the assumption that the concepts that lie behind a complex act are not all available at once", that assumption can be directly counterpointed with a correlated grammatical assumption, namely that the imposition of clausal structure upon a complex message works on the assumption that the speaker's cognitive apparatus allows him to have all the concepts available at once. The relationship between the two principles must therefore be an interface issue, with linearity as on the one hand an 'output constraint' (words have to come out in serial order) and on the other hand a pressure to minimize processing cost – the less you have to keep in the buffer, the simpler the production task.

But simplicity is not everything. The simplest way of moving from one act to the next in linear order is stimulus-response, the only format recognized by behaviourist psychology. And the evolutionary significance of grammar is precisely in establishing a level of communicative ability that defies the pressure to act directly, in a hand-to-mouth fashion, to immediate situational input and impulse.

6. A case for simultaneity of interaction, representation and expression in grammar: Choice of subject and fronted objects

Information structure is to some extent orthogonal to layered clause structure. Syntactic function assignment and constituent order can impose competing information structures on the same layered configuration. Because of this intra-grammatical interface effect, these phenomena are well suited to illustrate complexities of the procedural knowledge that speakers must have in order to make full use of the potential of the code they are using.

The most celebrated division of labour in the history of clause meaning is the division between subject and predicate. Roughly, it reserves a slot for the primary clausal topic (or, the constituent from whose perspective the clause meaning is viewed, or the constituent that is the prime candidate for full focal attention, cf. Tomlin 1995; the
precise difference is not important for this discussion), and presents the rest of the predication as information provided about it. In arguing that subjecthood is ‘purely structural’, a standard line of reasoning is that because it does not correspond to a specific semantic role such as ‘agency’, it cannot be semantically motivated. The mistake inherent in this reasoning is evident if you view the issue top-down, rather than bottom-up. If you view the subject as something carved out of the whole message, and something that must be selected in all ‘indicative’ clauses, it would be impossible to tie it down to the semantic role of agent, for the simple reason that agents are only available in clauses with an agentive verb. In contrast, the question of what is the most topical argument can be raised in all clauses consisting of verbs and arguments.

If the process basically works top-down, it also means that the choice cannot essentially depend on any other specific item-level properties. It belongs at a level of generality above all specific types of clause content; and this also explains why the role of ‘main clausal topic’ is not always equally relevant. If there is a general top-down coding strategy that must single out one argument as the subject, it is to be expected that it does not match all utterance functions equally well. As an extreme case, the existence of states-of-affairs without any constituent entities (the ‘atmospheric predicates, as in it is raining), means that there will be a clear mismatch between the high-level coding strategy and the lower-level filling material. In languages like English, the clash between an obligatory choice and the absence of semantic candidates results in an obligatory formal subject, i.e. it is resolved in favour of the high-level coding strategy.1

This, cf. Pinker (1994), can be used to argue for the lack of semantic content in syntactic categories. But that is a misunderstanding of the relation between (expression-)syntactic and semantic (=content-syntactic) complexity. In indicative clauses, there is a semantic (in the sense of encoded) choice between conveying a proposition as fact (declarative) and presenting it as subject to confirmation (interrogative). This choice, like the choice of most topical argument for subject, is bound up with the choice of the indicative sentence type in languages like English; in Danish, it also goes with the need to fill the ‘fundament field’, i.e. the pre-verbal slot, cf. below. Since these choices go together, it is functionally motivated to involve ‘subject position’ in encoding the declarative/interrogative choice. If clauses with zero-argument verbs had no subject expression, they would therefore not be able to reflect that choice in the same way as other indicative clauses. Thus, formal subjects (=subject expressions with no item-level content) are part of the functional-semantic organization of clauses, viewed top-down, even if they are empty from the bottom-up ‘item’ perspective.

This interface can only be captured if the model allows top-down and bottom-up approaches to the same overall procedural knowledge. Very roughly, if we follow the bottom-up approach, systematically getting the input in place before we assign higher

---

1. A related issue, which would take us beyond the scope of this paper, is presented by “thetic” constructions such as there is always Joe, where there is no clausal topic in the typical sense, and a NP is presented as information about ‘the situation as a whole’.
functions to it, the procedure will appear as a series of choices, ending up with the overall function of the whole thing. If we follow the top-down approach, selecting higher functions first, these will exert constraints with respect to elements that have not been provided yet by the procedure: if you have started off with clearly, you have to come up with a propositional content that you want to present as something obvious to your audience – a constraint that may prove communicatively suicidal.

In this section I am going to argue that the simultaneous availability of the whole procedure should include all three main levels, and that the element of online flow should be rigorously excluded from the domain of grammar, including expression. Although Bakker and Siewierska (2004) make a point of their model being temporally dynamic, nothing essential is lost as far as I can see if their model is interpreted as describing procedural knowledge rather than production. The same applies to Butler (this volume): The ‘dynamic interpretation’ of grammatical rules that he envisages can also, and in my view preferably, be understood as involving a demonstration of how the procedural knowledge embodied in the grammar can be used in getting from a conceptual structure to a linguistic utterance, rather than as a mechanical production mechanism. Butler’s discussion refers to the speaker’s “choices” (e.g. of what to put in the initial position) – and that would hardly be compatible with an understanding of the grammar as a production mechanism that mechanically converts conceptual input to linguistic expressions. In contrast, it fits well with an understanding of grammar as procedural knowledge which can be activated in different ways, depending on the actual circumstances (which include choices made by the speaker).

Under such an interpretation, the account suggested here is therefore fully compatible with dynamic implementations such as those suggested by Bakker & Siewierska and Butler. I agree also with Butler’s view that ‘psychological/cognitive adequacy’ must take processing into account: there must be a plausible path from the procedural knowledge to the production of actual sentences. The implementations suggested may even in some sense represent ‘canonical’ ways of going from knowledge to production, and as such constitute a separate and essential piece in the whole picture that includes both grammatical structure and discourse flow. The point I emphasize is that the speaker’s grammar exists also between utterances, and does not precisely characterize the way in which any concrete utterance is produced. If we take a given sentence and imagine a speaker using it several times (e.g., to explain why he arrived late), the successive production processes would be likely to change as a result of increasing routinization, and after a certain point it would probably be activated as a single chunk. These changes, however, would be irrelevant to a description of the grammatical structure of the clause (cf. also the discussion of Wray below, Section 8). The type of phenomenon I discuss is bound up both with interactive, representational and expression aspects of clausal structure, as well as being central in terms of incremental production.

If we regard grammar as procedural knowledge, standing apart from actual production, there is also a possibility of allowing actual processes of arriving at a complete utterance to move both top-down and bottom-up in relation to the same grammatical
hierarchy. The simultaneous availability of a top-down and bottom-up perspective is advantageous in the following way: On the one hand, status as clausal topic is an aspect of the intended communicative act (‘what am I going to talk about?’) which takes priority in relation to the details of clause content, and involves the central referential act that belongs at the interactive level in FDG. On the other hand, the assignment of subject status needs an input that cannot be specified without presupposing something about the representational structure of the clause: the subject has nuclear status in the predication that constitutes the core of the clause (non-nuclear constituents cannot be assigned to subject). And in assigning subject status one simultaneously decides on some expression properties of the clause, including template position options. From the point of view that I advocate, what is needed is an account where all these dimensions are explicitly linked. The procedural knowledge specifies in what way these three aspects are linked in one and the same choice, and whether you start with expression, predication or topic reference, the other dimensions should automatically ‘light up’, metaphorically speaking. The same thing applies (to take an example mentioned by Hengeveld 2004a), to the choice of a term or noun phrase in the slot after a linking (copula) verb. As he points out, it has the effect that the interactive act of ascription is realized by a syntactic constituent that is normally associated with referential functions – and that is captured by explicitly distinguishing between the interactive and the representational level. What is less clearly captured by the model is that putting a noun phrase rather than an adjective in the slot after ‘be’ is already a ‘rank-shifting’ choice at the level of syntactic composition. Not all verbs allow it (become crazy/become a nut; go crazy/*go a nut), and so it needs to be described simultaneously as an interactive choice and a choice at the core syntactic level.

There is nothing in the FDG format that prohibits this. But the whiff of ‘flow’ in the model makes it tempting to put such choices at one specific point in the procedure. In Bakker and Siewierska (2004) subject is thus regarded as something settled purely at the expression level, on the basis of properties assigned to the subject constituent as a ‘grammaticalized topic’. From the ‘flow’ point of view, that makes good sense. A production process would very likely function as Bakker and Siewierska suggest: subjechts are purely a consequence of properties assigned first at the interactive level and later at the representational level, and the expression properties that go with subjechts are simply assigned to the relevant constituent. However, if the facts outlined above, such as the availability of the choice between two versions of the same propositional content in Joe resembles Jack and Jack resembles Joe, are assumed to be part of the grammar of English, a procedure without a content-side choice correlated with the expression properties does not capture the functional-semantic dimension of subjecthood as part of the structure of languages like English.

The format I envisage would allow the top-down choice of ‘most topical argument’ to operate upon the output of the bottom-up process of generating the propositional structure, which provides the array of candidates for subjecthood. In grammatical structures which offer less than two candidates, subject choice will be ‘empty’ in the sense that there is no alternative, just an extra label to stick on the one available candi-
date. In the case of two- or three-place verbs, however, there is a choice (in languages like English). In this way, if subject assignment, with its functional content side, is part of the overall grammar, we can allow both for the existence of empty choices and for the possibility of actual production processes where high-level choices typically constrain the choice before an array of options has been lined up.

The same thing applies to the selection of P1 filler, the occupant of the first position in the clause. As in English, first position in Danish can be used for the subject as a kind of default option, or for a number of different other constituents, corresponding among other things to different ‘message modes’ as demonstrated by Hannay (1991). But there are a number of differences between the two languages, some of them plausibly associable with the fact that Danish is a strict V2 language, while English is predominantly SVO and only ‘restricted V2’ (cf. Vikner 1995). Roughly, the presence of an all-purpose preverbal slot, the ‘fundament field’, abbreviated FF (following Diderichsen 1946), in indicative clauses means that speakers have a more generalizable fronting option than speakers of English. As argued in Heltoft (1996), Danish appears to favour ‘linkage’ in filling FF while first position in English is more oriented towards ‘urgency’ (cf. also Harder & Poulsen 2001). The coding of ‘linkage’ entails, among other things, that pronominal objects occur readily in the Danish FF if they constitute the link with the previous clause, while this is less frequent in English.

As a way to investigate this difference empirically, Harder and Poulsen (2001), set up an experiment with a number of picture stories with four or five pictures showing a simple course of events. Briefly, the stories involved a main character who acted upon two similar objects, choosing the second over the first. One story could be briefly paraphrased ‘A baby gets a pear from its mother, but doesn’t want to eat it. Then the baby gets a banana, and he eats it’. In the cases discussed, there is no difference between English and Danish when it comes to subject choice – but there is a clear difference when it comes to filling the first position with direct objects. The result was (in the stories that ‘worked’) that more than fifty per cent of test subjects fronted the contrasted objects in the relevant place, whereas none of the English test subjects did. Literally translated into English, the Danish response was typically something like

(2) There’s a baby who is supposed to eat. His mum gives him a pear but THAT he doesn’t want. Then he gets a banana, and THAT he likes.

What the experiment showed in this context is that the language made a difference to fronting the objects, so the choice was not dictated purely by universal discourse properties of first position (such as have been demonstrated by Gernsbacher 1990). Subject and (to a slightly lesser extent) object were overwhelmingly the same in both languages. Filling the FF thus reflected a discourse property while simultaneously involving the expression component – but via the medium of the representational relationship between arguments in the predication. There is no plausible way to separate these dimensions in relation to the grammar.

In the paper we suggested an ‘editing-for-speaking’ effect (with a take on Slobin’s concept of ‘thinking-for-speaking’, cf. Slobin 1996), encouraged by a difference be-
tween test subjects who were shown the pictures one at a time and those who had time to familiarize themselves with the whole course of events: the latter fronted objects significantly more frequently. The options offered by the Danish FF apparently prompted an ‘editorial’ reorganization that was not equally available to English speakers (and there were no obvious ‘compensating’ strategies in the form of, e.g., intonational signalling of contrast).

One advantage of a theory that sees this as an interface between online cognitive processes and grammar viewed as a procedure, is that language-particular structure influences the choice of what goes first – although both grammars allow speakers to begin with objects (no absolute constraints prevent it). Thus the idea that the choice of first constituent occurs at a ‘purely cognitive level’ untainted by grammatical considerations is difficult to reconcile with these findings. Rather, the functional division of labour between the slots offered by the two grammars is different in the two languages – and thus speakers choose differently in the two languages.

Mackenzie (2004), following Levelt (1989), rightly points out that what has been talked of as ‘choices’ are automatized unconscious cognitive events. However, conscious awareness is not criterial if we see grammar as reflecting a socially available choice procedure; the availability of the possibility of putting alternative constituents in the FF, otherwise keeping the clause constant, is a fact about Danish which means that objectively there is a choice between different variants, no matter how this choice is actually executed. Also discourse choices, standardly viewed as fully conscious in contrast to grammatical choices, may be automatized. Scardamalia and Bereiter (1987) have shown that competent writers typically use a so-called ‘knowledge-transforming’ process, reorganizing their knowledge consciously so as to serve the purposes of writing an appropriate text, while less competent writers use a ‘knowledge-telling’ process, writing down knowledge as it reflected their current cognitive state, without being able to re-organize it in response to the writing task. However, Schilperoord (1996) has investigated a type of discourse production where true expertise was of the ‘knowledge-telling’ kind: attorneys dictating legal letters have a well-entrenched cognitive schema that allows them to dictate accomplished business letters without a moment’s reflection, because the knowledge automatically sediments itself into the relevant form. In the ideal case, no conscious choices have to be made, because everything goes to its proper place automatically, once the decision to dictate the whole letter has been made.

As suggested by Mads Poulsen (in prep.) there may be a cline between cases where the speaker blurts out what is in his mind, and is then ‘constrained’ to construct an utterance that allows that phrase to occupy FF position, and carefully crafted utterances where the coding options of the language have been optimally planned to reflect the speaker’s communicative intention. A case in point was when a Danish professor wrongly said that an author mentioned was a physicist, and then immediately corrected himself, saying

(3) matematiker er NN
    mathematician is NN
This constituent order is grammatically possible in Danish, but non-canonical in this context. Thus pressures of production may win over canonical content-syntactic procedures – but that can only be captured precisely if we have an account of the canonical choice procedure, and then account for ways in which the procedural knowledge specifications may be bypassed when they compete with other concerns. But a description of the grammar only as constraints would be misleading in terms of the ‘enabling’ role of grammatical options, as ways to provide whole coded utterances with the optimal functional properties, including discourse-functional meaning.2

7. Formulaic language

Formulaic sequences are interesting as a test case for the relation between grammatical knowledge and actual production, because they embody a critical duality that impinges on the production process: formulas are units, but they would not be special as units if they were not also internally complex in some way. Working from the perspective of corpus linguistics, Sinclair (1991:110) pinpointed the issue as involving two different coding styles, the ‘idiom principle’ and the ‘open choice’ principle, where the open choice principle is associated with a Chomsky-style view of creativity and grammatical compositionality, and the idiom principle is associated with choosing words in clusters that are already associated in the mind, or ‘prefabricated’. The open choice principle applies equally well to classical FG, while the idiom principle is closely associated with an aspect of FDG, namely the idea of letting certain phrases, such as ‘can I have…’ go directly from the interpersonal level to the expression component, because it constitutes a ready-made unit, cf. Mackenzie (2004). More specifically, the dual approach that is thus built into FDG is close to the position of Wray (who is cited in Mackenzie 1998). In the following I discuss the issue in relation to the position of Wray (2001).

Wray’s dual analysis retains the Chomskian analytical approach, but only as one half of the new picture. As expressed in the conclusion, “it is not only a question of knowing the words that go together into strings, but also of knowing the strings of words that go together” (Wray 2001:281). In her approach, the compositional-analytic and the formulaic-holistic approach constitute different processing modes which the language user can alternate between. There are two problems with her position that are also relevant to Mackenzie’s position.

The most central one, in relation to the question of psychological adequacy and in seeing grammar as reflecting actual production, is the purely individual-based view of language that this implies. Because Wray does not consider the level of the so-

2. The example is reminiscent of Hannay’s “reaction mode”, and it is natural to wonder to what extent that ‘message mode’ is grammaticalized in English and not in Danish – as opposed to it being basically a discourse type of event that occurs when urgency pressure is high enough.
cially recognized properties of the code, she has a basic definitional problem which
surfaces at the end of the book where she wants to draw the whole discussion together.
In the overall diagram (2001:263), she lists three different unit sizes: formulaic word
strings (strings of words stored and processed holistically), formulaic words (poly-
morphemic words stored and processed holistically), and morphemes (bound or free,
including monomorphemic words). But as her theory implies that there are individual
differences that cannot be inferred from the general form of these sequences, she adds
(2001:264):

> Unit size is just an explanatory device for the linguist who is trying to make
> sense of lexical storage in terms of standard language descriptions.(…) but this
> is for representational convenience only: The units are not intrinsically dis-
> crete. Any internal structure, including word breaks, is externally and secondarily
> imposed(…).

If you follow this non-discreteness to its logical conclusion, as Wray does (2001:264),
"the crux of the issue is that all units in the lexicon are effectively monomorphemic".
Formulaic sequences thus conform to the standard format for lexical entries: a lexi-
con lists the basic unanalysable elements, not compositionally complex forms. But this
would appear to run against very basic intuitions in cases such as quotations one has
learnt by heart (as an example, Wray mentions Hamlet’s soliloquy as part of an individ-
ual’s lexicon). In an extreme case, cf. Bradbury’s Fahrenheit 491, one might have learnt
a whole literary classic verbatim. Although from a processing point of view it is stored
as one single item, to say that Pride and Prejudice is monomorphemic to that speaker
makes clear the need to distinguish between properties of the code and properties of
the speaker’s cognitive storage system. The code per se does not possess one morpheme
expressing adequate understanding of Hamlet’s soliloquy or Pride and Prejudice.

The affinity with the direct path in Mackenzie from the interactional level to
expression is reflected when Wray (2001:132–134) suggests that separate and alternat-
ing mechanisms are involved, where the ‘specialization in social cognition’ goes with
formulaic language, whereas the ‘grammatical analysis module’ goes down to basic el-
ements. The affinity between ‘phatic communion’ and undifferentiated, holophrastic
signals on the one hand and propositional content and differentiated coding on the
other is obvious and undisputed. But for the purposes of a theory of grammar a gen-
eral affinity does not translate into an identification between structure and the purely
representational level of meaning, which would entail that interpersonal coding could
not take a delicately differentiated form. In connection with modal verbs in English,
for instance, choice between present and past tense typically encodes tentativity (may
suggest…/might I suggest…). Tense is a well-defined grammatical slot that prototyp-
ically encodes localization in time but which is sometimes recruited for interpersonal
purposes. An adequate functional theory of grammar should include the close rela-
tionship between categories used for alternatively representational and interpersonal
purposes, and not locate grammar on the sidelines in relation to interpersonal mean-
ing. Mackenzie (2004), in fact, provides a beautiful example of grammar being used
to provide differentiated encoding of interpersonal meaning: he shows how an expression that is parsable even if it is a formula (*can I have grapes?*) could be grammatically baroquely upholstered with purely interpersonal meaning, while the representational content remains the same (*If you could possibly see your way clear to providing me with some grapes...*).

One central exception is ‘reflexive’ utterances of the ‘ouch!’ and ‘damn!’ type (Wray 2001:256). Their special status is warranted by the fact that this class of utterances are represented subcortically in the brain, operating involuntarily in relation to relevant stimuli, and are thus the true analogues of animal calls, cf. above. Thus Hengeveld (2004) has a cognitive rationale for saying that an utterance like *damn!* bypasses the ordinary lexicon. But that is the only class in Wray’s inventory of formulaic utterances which has this property, and their status should not be extended to all ready-made utterances or (especially) fragments. A superficially similar but different case is when a formula of whatever size (including a single word) is used for a whole discourse act. In that case, the only process that is likely to be involved is item retrieval, with no structural formulation being involved – but that is a different matter from saying that the formula is a purely interactive, reflex-like response. A proverb like *it never rains but it pours* would appear to be a fusion of representational and interactive meaning, not a non-representational signal like *damn!*. And, pace Mackenzie (2004), the path through structural formulation is necessary when the formula invoked is not used as a complete act but has distributional properties that are grammatical in nature, such as *can I have*, which obligatorily takes an object and thus reflects one of Mackenzie’s prime examples of ‘declarative’ constraints, cf. the case of EAT as discussed above.

The danger of making the individual’s lexicon criterial can be anecdotally illustrated with a recent experience, where a student in charge of a meeting gave a short welcoming speech giving practical details such as the location of toilets, etc. It sounded perfectly spontaneous, but we happened to know that it had been learnt by heart (which is why it did not reflect information he had just received about the toilets being locked!). The whole string, in fact, was recalled as a chunk. This had taken some effort, and only gradually did the bits and pieces come together in a fluent presentation. A description of the production process would be very different from the first to the last performance, more and more levels of differentiation dropping out, until it functioned as one lexical recall. The difference, however, would hardly be relevant from the point of view of a description of grammar, or of the grammatical knowledge of the speaker. The quotation from P. G. Wodehouse below can demonstrate another facet of the issue, i.e. that even where there is no dispute about formulaic status in the lexicon, we cannot assume a straightforward relationship with actual production:

> ‘I noticed that the boy’s manner was sullen when I introduced him to Mr. Baxter, and said he was going to be his tutor. He disappeared into the shrubbery, and just now, as Mr. Baxter was standing on the drive, George shot him from behind a bush.’
‘Good!’ cried Lord Emsworth, then prudently added the word ‘gracious’. 

P. G. Wodehouse, Lord Emsworth and others (Penguin edition), p. 23

As illustrated, gracious can be an increment from the point of view of the cognitive production processes of the speaker, while in all other respects it is part of an unanalyisable whole unit.

The same point is relevant in relation to a construction-based view of grammatical structure. Constructions can be much more abstract than chunks, but are similar in that they emphasize the role of wholes over compositional processes. However, the importance of ‘gestalts’ does not, any more than the importance of formulas, detract from the freedom of human speakers to make more than one selection from the code in selecting what goes into a single communicative act. Whichever way you approach grammar, you cannot capture its unique contribution to human communication without allowing grammatical units a status different from that of discourse units.

8. Grammatical and other units in online communication

I am going briefly to mention how an interface approach between grammar and the flow of discourse can be enriched by other forms of segmentation, as demonstrated in Steen (2005) and Hannay and Kroon (2005). The central unit in this connection is the tone group, which Chafe (1994) has found to have many of the ‘increment’ properties that also play a role in Mackenzie (2004). Hannay and Kroon explore the interface between discourse and grammar with a view to identifying that unit of discourse which should be target in linking up with grammar, the ‘discourse act’. Based on a number of empirical examples, they argue convincingly that tone groups, corresponding in (English) writing to punctuation units, correlate closely with the steps of discourse progress understood as ‘strategic’ rather than conceptual acts. The word ‘prosodic units’ covers both types.

Discourse acts need not coincide with clauses, cf. below:

(5) He killed his wife. With a knife.
(6) I didn’t make you lose it what are you talking about!

Hannay and Kroon cite (5) from Ford et al. (2002), and follow them in describing it as constituting one grammatical clause while consisting of two discourse acts. (6) is from Bolinger (1989:97), who uses it to exemplify two clauses constituting one act. Both cases exemplify that segmentation into discourse steps is orthogonal to grammatical organization. Units of the flow of actual online action can be superimposed on the units of the code without being bound by them.

I concur in the analyses proposed, and my claim that grammar arose as a differentiation of the basic unit of communicative action from structured into unstructured items needs a rider in the light of that. The essence of having differentiated coding is to be able to escape the one-to-one relation between ‘interactive unit’ and ‘signal’, ren-
dering impossible a direct stimulus-response analysis of communicative activity. Once differentiated encoding of a single act is possible, the most obvious direct result is that single acts become sub-differentiated in terms of coding – but once the two levels are not mutually bound to each other, mapping between units of acts and coding units is also free to vary – as it could not before the big bang that splintered the holophrase. Thus in (5) clausal structure as a unit of compilation can be closed off after *wife* (but not before it), all dependency relations having been satisfied. When the next act constitutes a sub-clausal fragment in terms of coding, dependency relations require an interpretation whereby the instrumental meaning is attached to something in relation to which it can indicate the instrument. In other words the clause whose meaning had been completed (in terms of production as well as reception) is re-opened and re-recruited to complete the meaning of the new act. So perhaps a full description at the online level should strictly speaking be that there are two clauses involved – not two clauses corresponding to the punctuation units, but one corresponding to the first, and one corresponding to that reconstructed combination of the first and the second that is necessary in order to interpret the preposition phrase *with a knife*.

The semantic dependency relations that arise as a result of subdifferentiating message meanings into fragments can also throw light on another interesting example discussed by Hannay and Kroon, where the word *then* is given a whole paragraph to itself (heralding a disastrous turn of events). In terms of scope, initial *then* (at the top end of the bottom-up compilational process) is semantically designed to introduce a subsequent (course of) events, expressible as a possibly compound clause. When strategic segmentation into discourse acts severs it from the scopal content on which it depends, the effect is to create an expectation in the mind of the reader – suspending completion of the semantic compilation. The result of this suspension is – suspense: we know that something is missing, and we ask ourselves what it can possibly be. Thus the very fact that something is a semantic fragment can enhance its effect as a single strategic act.

Steen (2005), like Hannay and Kroon, pursues the issue of what unit is central for analysing the progress of discourse. Rather than focus on intonation units, he suggests that four different types of units are relevant (conceptual, prosodic, linguistic and communicative). The prototype unit he suggests is one where a proposition, a tone group, a clause and an illocution coincide. Both articles stress the need for further clarification of the relationship between the different dimensions. FDG may have a useful role in that process – but only if it avoids the danger of separating interactive, clausal and expression choices, and recognizes the special role of grammar as describing the units of the code rather than of the flow.

9. Conclusion

I have argued that the rise of grammar was a ‘big bang’ that splintered the holophrastic and unstructured ‘one-signal-at-a-time’ form of communication, and liberated
communicators from direct context-dependence. Thus grammar is fundamentally a complication, something that spoils the nice and simple relationship between encoding and interacting. Theories of the relationship between grammar and discourse must reflect this complication of the relationship.

The incremental perspective, however, still pervades the process of discourse, and thus fundamentally shapes the way grammar is used. Moreover, I argue that grammar, and grammatical knowledge, is fundamentally procedural, consisting in instructions for compiling complex clausal meanings – instead of constituting a rigid representational whole. Thus there is every reason to view grammar in the light of its contribution to the discourse process. However, the ability to compile complex meanings depends on the ability to handle all aspects of meaning as simultaneously available (regardless of what may happen in individual cases) – whereas the essence of online real-time production is the pressure to maintain linear flow. Thus we need to recognize the essential difference between flow and grammar in order to understand how they manage to collaborate. Having something called a ‘discourse grammar’ may wrongly suggest an identity between discourse constituents and grammatical constituents. That danger can be counteracted if the different levels are viewed in parallel, as simultaneously available aspects of a complex phenomenon. However, to the extent grammatical processes are seen as operating dynamically, in real time, it is tempting to understand them as distinct stages, reflecting a production process moving from communicative intention via formulation to expression. This brings about a risk of separating interactive, representational and expression phenomena that belong together in the code.

I have discussed some phenomena, especially subject assignment and FF-filling in Danish to illustrate the advantages of an account in terms of a grammar of procedures – keeping interactive, representational and expressive dimensions together in the same overall pattern – rather than actual discourse events. The basic locus for grammar, therefore, is not as an integrated element in actual production, but as something that can be invoked by the production process.

References

Grammar, flow and procedural knowledge


Heltoft, L. (1996). Topics in Danish and in Universal pragmatics. In L. Heltoft & H. Haberland (Eds.), Proceedings of the 13th Scandinavian Conference of Linguistics (pp. 251–262). Roskilde: Roskilde University, Department of Languages and Culture.


Grammar, flow and procedural knowledge


The non-linearity of speech production

Michael Fortescue
University of Copenhagen

The traditional linear conception of speech production shared by the majority of formalists and functionalists alike envisages a step-by-step succession of stages on the way from conceptual intention to speech chain product. However, the alternative conception of parallel distributed processing has for some time been gaining favour in such quarters. This contribution presents facts from a number of languages, including English, that support the idea that speech processing must be in large part non-linear, including at the lower levels involved in lexical access and phonological spell-out. Top-down contextual processes must constrain bottom-up generation, even when the relevant contextual trigger appears to lie further ahead in the linear speech chain. A Whiteheadian perspective is adopted to help gain insight into the non-linear processes involved in speech production and comprehension.

1. Morphophonological processing

The traditional linear conception of speech production shared by the majority of formalists and functionalists alike envisages a succession of stages on the way from conceptual intention to speech chain product, whereby ‘deeper’ representations are converted stage by stage into corresponding surface utterances. This picture is indeed perpetuated by many prominent psycholinguists, although the alternative conception of parallel distributed processing has for some time been gaining ground. The explicitly non-linear models of the connectionists are still by and large limited to comprehension (mainly at the morphophonological word level) and, despite the opacity of their ‘hidden layers’, presuppose step-by-step processes whereby outputs come gradually to match input and ‘known’ target types. The new perspective on the nature of processes provided by non-linear dynamics and what is popularly called ‘complexity theory’ has in recent years gradually extended from analyses of self-organizing biochemical systems with their multiple feedback loops and massive redundancy also to apply to social processes (cf. Luhman 1990), yet this perspective still meets resistance in mainstream linguistics, presumably because of the still pervasive metaphor of communication as the exchange of information through a linear channel, in effect allowing only one phoneme at a time to pass through it. This is in fact a highly debatable con-
attention, and the alternative perspective, whereby linguistic communication is seen as a matter of mutual coordination of behaviour through linguistically coded clues, a kind of ‘structural coupling’ of speaker and hearer (cf. Maturana & Varela 1987:257ff.), is becoming more and more widespread among cognitively oriented linguists.

No one seriously doubts that the linearization of complex, non-linear intentions through the limited channel afforded by the speech organs must take place somehow (or to some degree). As Bates and MacWhinney (1987:211) put it: “…surface forms are multiply-determined. They are emergent solutions to the problem of communicating non-linear meanings onto a linear speech channel”. However, there is considerable difference of opinion as to the number of discrete stages (if any) prior to output – in other words, as to whether linearization is a one-off, real-time process (as implied by the production interpretation of Simon Dik’s functional model, for instance) or whether the resulting speech chain is just the last in a series of transformations or re-representations on the way from module to module towards final articulation (as implied by generative models inspired by Chomsky). Bates and MacWhinney’s own ‘competition model’ is radically emergentist and non-deterministic, but of course that much harder to formalize. At the highest level of ‘top-down’ planning, one might say that any model of speech production is by necessity holistic and non-linear, but at lower levels – namely those of morphophonological elaboration and phonetic ‘realization’ – one might a priori expect motivation for linear thinking to be more compelling.¹

What I shall be proposing in this paper is that even at the morphophonological level this search for pure linearity is illusory and that the Whiteheadian framework I espouse provides a uniquely fruitful perspective on the relationship between top-down and bottom-up aspects of speech production since it presupposes non-linearity from start to finish. Space precludes a detailed discussion of the pivotal notion of concrescence, so the following characterization will have to suffice (see Fortescue 2001:126f. for further discussion). The basic elements in Whitehead’s philosophy of organism are actual occasions. These ‘drops of experience’ come together from the data of their immediate past then ‘perish’ at once to become objective data for subsequent occasions. They consist of processes of self-organizing concrescence, whereby the inherited data – e.g. linguistic input – is grasped by prehensions (or feelings) and successively integrated according to a subjective aim (e.g. to comprehend that input) This aim is partly determined by the occasion’s perspective on relevant eternal objects (concepts, forms of definiteness, including grammatical patterns) and partly by the nature of the data itself. It strives towards the satisfaction of the occasion (the achievement of maximal

¹. Though even here autosegmental ‘prosodies’ belie the possibility of a purely linear approach. In fact, the whole idea of the single-channel nature of communication is highly questionable (hence the ridicule to which the ‘conduit’ metaphor of speech has been put by cognitive linguists). One need only consider the gestures and facial expressions accompanying speech, not to mention the radically expanded possibilities for simultaneous and overlapping signing provided by sign languages.
unity and intensity of integration of its prehensions) and grows more differentiated and explicit as the \textit{concrescence} proceeds. Actual occasions are organized into interrelated \textit{nexus} or societies. Persons, enduring objects, events and human societies are all types of \textit{nexus}. There are societies within societies within societies (all functioning simultaneously but at different levels), from atoms through cells to living bodies and beyond in an emergent relationship. The mind/brain is thus a hierarchically organized ‘society of societies’. In general terms, quoting Sherburne (1966:7): ‘Concrescence’ is the name for the process in which the universe of many things acquires an individual unity in a determinate relegation of each item of the ‘many’ to its subordination in the constitution of the novel ‘one’. An actual occasion is nothing but the unity to be ascribed to a particular instance of \textit{concrescence}. This \textit{concrescence} is thus nothing else than the ‘real internal constitution’ of the actual occasion in question.

Levelt’s influential production model (Levelt 1989), which incorporates Kemper’s notion of the ‘incremental’ approach to processing, combines parallel and serial processing in a way that is largely compatible with the Whiteheadian framework. It is also germane to Mackenzie’s ‘incremental’ approach to Functional Grammar (Mackenzie 1998 and 2000). Levelt’s lexical-based approach divides speech production into successive stages starting from a high level intention via the activation of relevant ‘lemmas’ (the meanings of lexical words and their morphosyntactic frames, minus their phonological form) in the ‘Conceptualizer’ component or module, then morphosyntactic encoding in the ‘Formulator’ (as constrained by the limited combinatorial possibility of each lemma involved), and finally phonological encoding in the ‘Articulator’. This model tries to reduce ‘look ahead’ processes to the minimum compatible with experimental and speech error data – an example he gives of where it \textit{is} necessary is in assigning stress to ‘double-stressed’ words like \textit{sixteen} in English, which depends on the prosodic structure of immediately following segments, thus \textit{sixteen mén} as opposed to \textit{sweet sixtéen}. The model is ‘modularized’ (automaticized) only to the degree that this appears to be warranted by available data on interference – or the lack of it – between components, and as far as its grammatical component is concerned is closely allied to Lexical Functional Grammar.\footnote{Contrast this with Hawkins’ (1994) Early Immediate Constitent parsing model, which is intended as a general framework for explaining morphosyntactic typological universals (in particular those concerning the interaction of word order and constituent ‘ heaviness ’). This is both encapsulated (completely independent of semantic input) and serial: it allows for no forward guessing about high-level structural organization at all. The parsing principles involved (performance) are seen as innate and interacting with equally innate universal principles of grammar (competence).}

The procedural algorithms involved are envisaged as unfolding essentially ‘from left to right’, starting with the most salient ‘message fragment’ in the Conceptualizer, but with the possibility of holding the products of lower-level constituent generation in syntactic and prosodic buffers while higher level nodes are elaborated (thus the generation of syntactic patterns around the major constituents of a sentence may proceed...}
in parallel). Moreover, the input to individual components may be multiple – thus the input to the Prosody generator includes both accreting syntactic ‘surface structures’ and ‘intonational meaning’ involving numerous parameters of an emotional and interactional kind affecting pitch, contour, speed of delivery, etc., that somehow have to be integrated into a unitary (linear) output. It is not spelt out, however, what exactly pre-linguistic ‘message fragments’ are – on the model they are merely represented by capitalized labels; in fact the whole model is of a decidedly ‘representationalist’ nature.3 Within the Whiteheadian framework such message fragments correspond to ‘prehensions’, i.e. to feelings rather than representations, guided by overall ‘subjective aims’. These hold out the promise of getting closer to the psychological ‘reality’ behind the eclectic AI-like metalanguage in which Levelt’s model is formulated. The explicitness of Levelt’s descriptive metalanguage (and that of other modelers of speech production) is of course essential for defining the areas and dimensions of the model to which well established psycholinguistic testing procedures can be applied, but it does cry out for phenomenological interpretation of the type provided by the Whiteheadian framework.

When one looks closely at languages with rich morphologies (and complex morphophonological processes) it becomes apparent that even at the largely automated level of phonological adjustment the involvement of non-linear processing is inescapable. Take the phenomenon of bidirectional vowel harmony in Chukchi: all vowels in that language (apart from schwa) are either dominant or recessive, and if there is one dominant vowel anywhere in a complex word then the vowels of the whole word must be dominant. This means that when a dominant suffix is added at the end of a long word-form otherwise containing only recessive vowels, all the preceding vowels must be changed to match. Thus the allative case form of recessive base *ngewtum*- ‘girlfriend, wife’ is *ngawtomgety*, where the vowels of the stem have been ‘replaced’ by the corresponding dominant ones because of the dominant inflectional morpheme. It

3. The role of static representations in a dynamic production model is perhaps the only serious difference between my own approach and that of Butler (this volume). Butler attempts to spell out explicitly all essential sub-lexical semantic factors involved in the mappings concerned. A ‘concrescence’ approach to the production of whole utterances would of course have to somehow take into account all these factors too (‘prehended’ in semantic memory as necessary), but it does not assume that they are necessarily involved in specific on-line processes. This is essentially a matter of emphasis: I am just a little less sanguine about accepting the direct reflection of the constituents and patterns of linguistic pattern in the processes of language production and comprehension. I agree fully with Butler that much remains for future research – especially perhaps as regards the integration of constraints on linguistic productions deriving on the one hand from grammatical structure as such and, on the other, from higher-order context and intention. No form of linguistic modelling has as yet been able to show how exactly such factors might interact ‘on line’. Both Butler and Levelt (and indeed Mackenzie) take as the starting point for their incremental models the most salient part of the input ‘message’, which must surely be on the right track, but we still know very little about the binding between ‘saliency’ and subjective perspective on linguistic ends and means available.
seems intuitively unlikely that the Chukchi speaker at some late stage in the production of this word has to move backwards through the speech chain in order to readjust the vowels accordingly. One could also mention here umlaut phenomena in more familiar Germanic languages (thus Icelandic nominative kött–ur 'cat', displaying u-umlaut, but genitive katt–ar, and plural kett–ir displaying i-umlaut), but this is much more lexicalized and the forms presumably learnt as wholes.

Other phenomena that immediately suggest a non-linear or parallel approach include Finnish 'consonant gradation', whereby a consonant closing a syllable determines that a 'strong' consonant or cluster at its onset is reduced to the corresponding weak one (as in avun, genitive of apu 'aid'). This is also apparent in connection with exceptions to the basic rule, such as the illative case form apaun (where the following vowel, being long, blocks the process), or possessed forms like kätensä 'his hand' (with personal possessor suffix -nsä not causing weakening – compare kädet 'hands'). It is even more apparent in cases where weakening occurs in unexpected environments such as infinitive putota 'to fall' from stem putota- plus infinitive suffix -ta, which actually reflects a two stage reduction process via an intermediate syncopated *putotta (the infinitive form actually ends in a historical consonant lost in absolute final position, and this causes the reduction of the preceding /tt/ to /t/, but only after the first consonant of the unreduced /tt/ has caused the preceding single /t/ to reduce to /d/). These 'look ahead' conditions may be morphologically as well as phonologically determined.

Of course many phonological and morphophonological processes found in more analytic languages like English (e.g. stress assignment, as mentioned above) or Chinese (e.g. the realization of the Mandarin low third tone depending on the tone of the following syllable) also involve looking ahead (or 'backtracking') and/or 'instantaneous' template matching. One might add that languages displaying switch reference, whereby the marking of 'same subject' or 'different subject' on medial verbs presupposes access to the subject of the following main clause, specifically argue for the necessity of syntactic 'look-ahead' processes. Linguists concerned with formalizing such matters have, at least since the beginning of the generativist era with Chomsky and Halle, taken it for granted that language is hierarchically (non-linearly) organized. And at least since Zellig Harris, awareness of the discontinuous aspects of English grammar (e.g. in 'phrasal verbs') has been center-stage in syntactic thinking, also as regards recent computer simulations of the production and comprehension of English sentences. Some have gone further and claimed that the actual linear deployment of speech is 'merely' a matter of performance. This I do not intend to argue with, but would suggest that what needs to be carefully distinguished here is not so much competence and performance as the more general perspectives of language-as-pattern and language-as-process, the confounding of which has led to much hand-waving in discussions of the interface between grammar and speech production. Taking abstract generative models in particular as somehow cognitively/psychologically real (as Chomsky would have it) – as opposed to maximally economical in theory-internal terms – may lead to cognitively absurd conclusions. For instance, Rice has proposed that the positioning of the stem consistently at the end of verb-forms in Athabaskan languages despite their
(S)OV syntax (otherwise inexplicable within the generative framework she works in) must reflect an underlying movement of the stem from initial position to the end of every verb (Rice 2000). Is this supposed to mean that Athabaskan speakers ‘in performance’ actually have to start at the end and reverse the ‘underlying’ order of the stem vis-à-vis all other morphemes in every verb they utter? If the movement is already taken care of by competence, then competence is indeed involved in the linearization process (which is otherwise assumed to be a matter for performance).

What I am moving towards is a proposal as to why speech production has to be non-linear, from start to finish. I am presupposing here a one-stage, holistically constrained incremental production model, one specifically based on the Whiteheadian notion of the concrescence, but also close to McNeill’s concept of the ‘growth point’, in which language is a matter of static templates or filters constraining dynamic pre-linguistic expressive intentions (McNeill 1992). No intervening representations between input and output are required. The patterns and constructions accessed in realizing an utterance as both grammatically well-formed and matching semantic/pragmatic intentions are from this viewpoint Whiteheadian ‘eternal objects’ (predispositions derived from – and reinforced by – experience) that are evoked to constrain the expressive process (cf. Whitehead 1978:22f.). There are language-specific eternal objects (or figurae, if one prefers) relevant to each level of linguistic structure, from the phonemic to the syntactic and beyond (their abstract generalization to Universal Grammar will not concern us here). Where there is potential conflict between such patterns, competition will result in the best ad hoc approximation to harmony in the context of the conflicting input – essentially an ‘aesthetic’ matter in Whitehead’s broad sense of that word. The speech chain is seen as emerging over a finite stretch of real time, starting from the beginning, but reaching ‘completion’ only as the process proceeds: the end need not be predetermined when the beginning is already being produced. No complete intermediate ‘representation’ at all is envisaged before actual production of the speech chain. The entire utterance is produced just as soon as it can be – and ‘suspension’ of the realization of the overall pattern while subordinate ones are attended to is an integral part of the picture. This is, in other words, a ‘wheels-within-wheels’ approach to speech production. The input is a teleologically dynamic, contextualized bundle of ‘prehended’ data that need to be mutually organized according to accepted patterns of expression (through eliminations and highlightings) in order to produce a well-formed output, an utterance.

One might argue that there is an inherent problem with any such single stage model, namely that even if it does correctly capture ‘psychological reality’, its hypothetical single process of linearization can be broken down for analytical purposes into successive transformations or representations, and that, given our present crude level of ability to localize specific processes within the brain, its ability to predict testable outward manifestation in speech production would be no different from that of the corresponding multi-stage model. Though I do not of course have privileged access to psycholinguistic data that could prove such a model today, I at least have Occam’s razor on my side: the model presupposes the logically minimal number of ‘trans-
formations’ or stages between intentional input and expressive output and does not hypostatize analytically convenient intermediate stages as ‘real’. Breaking a single dynamic process down into successive discrete steps is surely a descriptive illusion if some of these steps themselves turn out to be non-linear too – as is arguably the case in for example the conversion of an underlying phonological template into one marked for degrees of stress in English. The onus is surely on those who still favour multi-stage, multi-representational linearization models to show that the greater complexity of their models is actually justified by the accreting empirical data from psycholinguistic investigations.

2. The role of prehensions in speech production and comprehension

So what exactly might the non-linear processes involved in speech production and comprehension be like, and what is their ontological/physiological basis? This is where the Whiteheadian framework comes into its own, since it provides the means – the theory of prehensions – for directly linking speech processing with experiential ontology, something both psycholinguists and practitioners of AI tend to shy away from. The following is a summary of the basic types of Whiteheadian prehension that may be involved in speech production (compare the schema in Sherburne 1966:40). For a detailed analysis of a specific dialogue text in terms of these prehensions see Fortescue (2003), where the role of memory and attention span is also crucial.

(a) Physical

From the afferent (comprehension) perspective, this simple type of prehension – typical of the early ‘conformal’ stages of a *concrescence* – is involved in the analysis of the raw speech ‘data’ (phonological, prosodic, gestural accompaniments, etc.) and of relevant perceptual aspects of the communicative setting. Such conformal prehensions occur automatically at the neural level of the speech organs, not at the level of the conscious ‘personally organized nexus’ that concerns us here. For Whitehead, consciousness is a ‘pulsed’ kind of subjective form associated with certain types of (higher order) prehensions within personally organized nexus (read ‘organisms’ or ‘self-organizing systems’). Compare Damasio (2000:176) on the pulsed nature of ‘core consciousness’, each pulse (or Whiteheadian ‘actual occasion’) being paced by each new ‘object’ (external or internal) impinging on and changing the organism’s ‘proto-self’, thereby creating a pre-linguistic ‘narrative stream’ of interaction with its environment. The result of such a self/object integration is in turn new input for a further interaction. Beyond ‘core consciousness’ lies extended or ‘autobiographical’ consciousness (Damasio, by the way, acknowledges Whitehead’s influence in his book). The temporal pacing of the complex integrations between deep somatosensory ‘core consciousness’ and extended cortical memory may involve the thalamus as well as frontal ‘working memory’ loops (the thalamus may map object/organism relations in ‘implicit’ form on the way to more explicit imagery elsewhere – cf. Damasio op. cit.:221 and 265).
during objects – including experiential ‘mental models’ in memory – is also generally physical. Physical prehension of the speech signal in the mode of ‘causal efficacy’ leads via abstraction from the input data (by negative prehensions that isolate the signal from ‘noise’ and allow for higher-level contextual variation) to the recognition of sequences of individual phonemes in the more fine-grained mode of ‘presentational immediacy’, which constitutes a bridge to the following, ‘supplementary’ phase of the concrescence.

(b) Conceptual
This kind of prehension is involved in ensuing phases of the concrescence where specific concepts are recognized as being associated with the speech input chain (via phonological form and symbolic reference links). From the efferent point of view of the speaker, corresponding eternal objects (‘concepts’ in the widest sense of recurring patterns) are prehended as potentially contributing to the satisfaction of the initial subjective aim (the overall discourse goal). The lexical items evoked (along with their respective argument frames, etc.) are then integrated according to mutual compatibility with the general morphosyntactic patterns associated with the complex discourse act aimed at. Negative prehensions are needed in order to select for expression only the most suitable items thus prehended (see below). In this way the subjective aim of the concrescence (inherited from the speaker’s immediate past) becomes more and more determinate and specific through successive conceptual and propositional prehensions.

(c) Hybrid
Hybrid prehensions are the type involved in recognizing the intention behind another person’s utterances (by ‘empathy’, etc.). They are also involved in accessing relevant presuppositions and making necessary ‘bridging inferences’ in order to fill in inconsistencies in accruing ‘mental models’. They may draw upon either general principles or specific knowledge (the enduring objects of memory).

(d) Indicative
This is the kind of physical prehension involved in acts of reference (to a logical subject – including those ‘labelled’ by proper names), or of establishing a discourse topic (if there is a choice of different logical subjects). Such a prehension will highlight certain regions of the relevant nexus (here read ‘contextual network of relations’) for potential comment, etc.

(e) Propositional
This kind of prehension combines types (b) and (d) above, in the simplest case assigning a predicate to a singular logical subject. The ‘imaginative’ propositional sub-type
assigns a predicate to a logical subject in a relationship that is imagined rather than perceived. A proposition as such, note, is neither an ‘eternal object’ (a repeatable pattern) nor a nexus, but a hybrid relationship between logical subjects prehended within a nexus (as under (e)) and a predicate (an eternal object) of a kind that the human mind/brain is especially attuned to prehend. It is not necessarily expressed or articulated verbally. What is prehended (as with all higher level prehensions) is a species of contrast (between the nexus and the eternal object).

(f) Intuitive judgments
This kind of complex prehension is a combination of (d) and (e) at a still higher level of complexity within the supplementary phase of a concrescence, where a whole proposition is grasped within its overall nexus (thus allowing the assignment – or suspension – of a truth value). Differing ‘subjective form’ here may reflect varying epistemic stances on the part of the speaker. The intuitive judgment forms the basis of processes of inference and abduction of various kinds.

(g) Anticipatory
These ‘feed-forward’ prehensions monitor the progress of the concrescence towards the intended satisfaction of its initial subjective aim, in particular regarding the effect on the hearer that the final utterance can be expected to have (this may involve general principles of communication or specific knowledge concerning the recipient and the given situation of context). The later stages of the specification of the ‘subjective aim’ involve anticipatory prehensions.

There is also the ‘negative prehension’ (as already referred to under (a) and (b)), but this is best seen in the present context as a coverall term for the application of constraints of all kinds in processes of abstraction. They function in conjunction with the positive prehensions involved to heighten the effect of the latter and to resolve incompatibilities between competing prehensions in fulfilling the subjective aim. They represent the suppressing role exerted by those eternal objects that might have been prehended in the input data but are not relevant to the subjective aim (in doing so they enhance the contrastivity of those ‘adjacent’ eternal objects that are indeed relevant). One may state that every positive prehension – e.g. of a word – is associated with a negative one heightening its contrast with ‘what it is not’ (much as individual neural columns in the cortex are believed to function).

Other, more complex Whiteheadian processes include Transmutation and Symbolic Reference. Transmutation is the transformation of a series of physical prehensions (e.g. of an ‘enduring’ object) associated with a single eternal object via a conceptual prehension, so that it is abstracted as a unitary ‘thing’ contrasting with (and representing) the eternal object itself, the pattern type. Symbolic Reference is perception in the ‘mixed mode’ of presentational immediacy and causal efficacy, linking sensory signs (in the former mode) and their ‘felt’ meaning (in the latter), i.e. not a single type of prehension but a particular linkage between physical and other types of prehension. It is the source not only of linguistic/semiotic behaviour as such, but also
As a concrete exercise in the translation of specific cognitive processes into Whiteheadian terms let us consider the nature of lexical access, a fairly low-level (and therefore relatively well understood) stage in linguistic processing. I shall take it that some form of the bottom-up Cohort Model (Marslen-Wilson 1987) represents a close approximation to reality here – although it needs to be supplemented by top-down considerations of context for constraining the overall process. This the Whiteheadian framework can provide since it combines ‘on-line’ linearity of processes with higher-level holistic control. The basic idea of the Cohort Model is that lexical access is a linear matter, with the very first phoneme of the input string (corresponding to the word to be accessed) activating all phonological entries in the mental lexicon beginning with that phoneme; thereafter the number of potential candidates is rapidly narrowed down as each successive phoneme is analysed. At some cut-off point there is only one candidate left activated, and its associated semantico-syntactic ‘lemma’ (to use Levelt’s term) may in turn be activated for further processing.

The whole process can be treated as a single Whiteheadian concrescence (albeit one of comprehension rather than production), bound together by the common subjective aim of accessing the relevant word in lexical memory, which is satisfied when that aim is realized. This concrescence specifically involves the nexus pertaining between the acoustic signal and the phonological representation of the word in the memory of the hearer. The rapid and automaticized registration of a ‘match’ here may well always be below the threshold of consciousness relevant to higher level linguistic processing. It would seem a priori to fall largely within the initial ‘conformal’ phase of the concrescence, with only the final activation of the word’s lemma involving a higher ‘conceptual feeling’ (of the complex eternal object corresponding to the word’s ‘meaning’). However, as we shall see, conceptual feelings are in fact involved almost from the start. If we take the English word preference spoken in isolation as the target, we can schematize the process as in (1a) below as a linked series of stages; in (1b) we have a slightly more

6. All these different kinds of process must in some manner be ‘anchored’ in the brain, of course. All except the most basic physical prehensions (in the primary sensory and motor zones) presumably involve some activity in the (pre-)frontal lobes, the hierarchically most superordinate regions of the cortex where heuristics for the satisfaction of subjective aims are worked out. The ‘causal efficacy’ and ‘presentational immediacy’ poles of the relationship of ‘symbolic reference’ may reflect a cortical-subcortical distinction. Episodic – or autobiographical – memory, anchored in the hippocampus, is for Whitehead is at all events ‘perception in the mode of causal efficacy’ (Whitehead 1978:120), while ‘semantic’ memory, distributed in association areas of the cortex, would seem to involve primarily perception in the mode of sensory ‘presentational immediacy’ more directly. The skewing that favours the development of language skills in the anatomically more fine-grained left cortical hemisphere is another matter. This also favours holistic processing in the right hemisphere, e.g. in relating linguistic comprehension and production to broader shifting perspectives or contexts, but either hemisphere would appear to be capable of functioning to some degree in the joint mode of symbolic reference.
complex case involving a prefixed form that may well be a neologism for some readers. Here the prefix dis- is first activated, after which, as soon as the following syllable has been phonologically analyzed, the continuing analysis of the whole string as a single stem is aborted and a new search started for the stem that could follow the prefix initiated. The two morphemes are combined in the final stage. The prosody also helps the process along, since the secondary stress on the prefix undermines any attempt to find a fully lexicalized stem with secondary stress on dis- followed by primary stress on the following syllable (the secondary stress is thus an important clue to the presence of the prefix). Actual lexical 'entries' corresponding to individual morphemes are indicated in capitals and the point of unique lexical 'decision' shown by double arrows.

Note the parallel analysis of the overall prosodic package from raw 'analogue' signal to 'digital' syllable structure (much simplified).

<table>
<thead>
<tr>
<th>phonemic analysis</th>
<th>prosodic pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1a) 'preference'</td>
<td></td>
</tr>
<tr>
<td>pro……………….</td>
<td></td>
</tr>
<tr>
<td>pr……………….</td>
<td></td>
</tr>
<tr>
<td>pr……………….</td>
<td></td>
</tr>
<tr>
<td>pr……………….</td>
<td></td>
</tr>
<tr>
<td>pr……………….</td>
<td></td>
</tr>
<tr>
<td>pr……………….</td>
<td></td>
</tr>
<tr>
<td>pr……………….</td>
<td></td>
</tr>
<tr>
<td>pr……………….</td>
<td></td>
</tr>
<tr>
<td>pref…………….</td>
<td>/</td>
</tr>
<tr>
<td>pres……………</td>
<td>___________</td>
</tr>
<tr>
<td>prev……………</td>
<td></td>
</tr>
<tr>
<td>prefas…………</td>
<td>/  \</td>
</tr>
<tr>
<td>prefarons………</td>
<td>__ __ __________</td>
</tr>
<tr>
<td>PREFERENCE</td>
<td></td>
</tr>
</tbody>
</table>

7. It was for me when I first ran across it in Hall (1992), though the context – a discussion of the typological preference for suffixes over prefixes, in which the converse (the 'dispreference' for prefixes) was a natural inference. Hall's point was to show that in cases of 'flirting morphemes' where affixes are wavering on the point of forming new lexicalized wholes, there is greater resistance to prefixes than to suffixes because the former requires more complex processing, namely two 'cohorts' rather than a single one (the second – accessing the stem from scratch – starting only when the prefixed morpheme is accessed). This is reflected in (1b).
The first phase of the *concrescence* is the abstraction of the task-relevant 'objective data' from the raw signal. One might take the Cohort Model literally and say that this is constituted by the abstraction of the phoneme /d/ from the onset of the signal, then the phoneme /i/, and so forth, step by step. But already at that stage we must surely be dealing with parallel processing in so far as not only a chain of phonemes but also the construction of a tiered syllabic structure to match an overall prosodic template has been initiated. And if the 'objective data' is only extracted from the signal in step-wise fashion can we talk of a unitary *concrescence* at all? I believe so: it is the unique subjective aim that determines this. Within a given *concrescence* one cannot talk of a strictly linear temporal sequence – this must remain indeterminate – but only of a *logical sequence* actually carried out in some sense 'all at once', as constrained by the real time required for the functioning of the different sub-processes drawn into its 'satisfaction'. The principal temporal constraint here is the 'linear' nature of the acoustic channel itself.
Now a single Whiteheadian *concrescence* is constituted by numerous individual (if overlapping) prehensions. What are the prehensions involved in this particular task? They include at least some simple physical prehensions, but also (at a logically subsequent stage) some simple conceptual prehensions (recognizing relevant patterns in the data). Moreover, every time a new phoneme is ‘recognized’, negative prehensions suppress all candidates that are now eliminated. A simple physical feeling (or ‘prehension’) is one that feels another feeling – in this case the ‘feel’ of the input data is recognized by the auditory system for what it is: a speech signal. That is in fact the initial step of the *concrescence* (the ‘conformal’ stage, which merely causes the receptive agent (the brain) to isolate the signal from surrounding ‘noise’ as it prepares for more specific analytic processes, namely ascertaining the chain of individual phonemes the signal contains. This is in turn achieved through conceptual prehensions of the individual phoneme patterns in the hearer’s linguistic memory (their readiness for potential reaction) whereby each phoneme is experienced as ‘the same’ as the feeling of the section of the objective data corresponding to it, i.e. there is a ‘match’.

What then of the status of the individual ‘conceptual’ patterns thus abstracted, the individual phonemes and the higher level (syllabic) organization of the word that is accessed? They are themselves all ‘eternal objects’ relevant to different levels of linguistic organization, patterns somehow instantiated in memory that are ‘permanent’, or rather repeatable, with a stable if context-sensitive value. It can be seen here why simply translating ‘eternal objects’ as ‘concepts’ can be quite misleading, since the term also covers repeatable perceptual patterns.

What the Whiteheadian perspective can add to the Cohort Model is, as I have claimed above, the lacking top-down dimension. Beyond the ‘bottom-up’ data provided by the input signal, the activated candidate phonological words are also contributing to the overall process by anticipating their successful ‘closure’ (through ‘anticipatory prehensions’). Moreover, at a higher level, their associated lemmas are also anticipating their closure in unique activation, and still broader contextual expectations are generating further anticipatory prehensions. In so far as unique access to the lexical item concerned does lead to further processing involving the lemma of that item, we may talk of a higher level *concrescence,* the aim of which is to search for the meaning of a whole utterance-in-context. This overarching *concrescence* (as sustained or ‘entertained’ by both discourse participants) can be said to prehend the lower level *concrescence* as part of its own objective data. The process is constrained by pathways of causal efficacy leading from the past to the future for each of the speaker/hearers concerned. The closure (or ‘satisfaction’) of the uniquely targeted phonological word is in fact constituted by the activation of its anticipated lemma – and the closure of *that* is constituted by its ‘fit’ within a still broader context of speech.8 The resultant

8. See Pulvermüller (2002) for some hypothetical neural mechanisms within a distributed parallel processing paradigm whereby sequential and ‘pushdown’ hierarchical relationships between words within whole sentences could be recognized (by “sequence detector” aggregates and multiple resonance states of “functional neural webs” respectively). The “closure” achieved
'objectivized' feelings get passed on (upwards) from level to level as the whole system seeks closure, i.e. comprehension. Closure at the highest level may ultimately involve the right hemisphere, which can function as an ‘anomaly’ detector between on-going linguistic (as well as sensory) activity and relevant mental models (cf. Ramachandran & Blakeslee 1999:280, Note 9).

It may be that fully activating the lemmas corresponding to individual words is only necessary when there is a lack of immediately ‘felt’ conformity between the input and the model of the on-going discourse (notably when new or contradictory meaning is involved, or when its pragmatic motivation is obscure). Lemmas may be ‘primed’ through various channels, including phonological form and partial (context-determined) aspects of the overall meaning potential they cover. Where full analysis of the meaning of an individual word in the input is essential this can be viewed as carried out through a concrescent process of parallel feed-back and feed-forward between phonological form and overall context, mediated by the word’s lemma. The ‘relevant contrast’ which the word symbolizes thus becomes homed in on through successive prehensions until a satisfactory (if approximate) match of content to context is attained and concomitant adjustments to the model initiated (e.g. by adding a proposition to it). Note how this proposal combines linearity of expression with non-linearity of overall processing (within individual concrescences), which is precisely what Levelt’s model strives towards, and which, as I understand it, also lies behind Mackenzie’s notion of Incremental FG: for from that perspective what unfolds in linear fashion from the very beginning of the utterance contains, so to speak, the whole in ovo.

I hope that the preceding account will have convinced the reader that the Whiteheadian framework can indeed complement other approaches to speech production in a useful manner, despite its initial terminological ‘foreignness’. What I have attempted to illustrate is how the theory of the intention-driven concrescence and its component prehensions – especially perhaps those of the anticipatory variety – can provide a top-down, holistic perspective that can not only help clarify the non-linear aspects of speech production but also explain their necessity, aligning speech production with all other modes of mental process. Without the constraints provided by high level subjective aims, low-level processing would ‘explode’ along all manner of conceivable alternative pathways, like a Cohort Model run in reverse. Unlike other approaches, the one I have sketched in this paper combines consideration of the production of linear speech chains with consideration of the multi-dimensional intentions that may modulate such processes en route from start to finish. The black box of Levelt’s ‘Conceptualizer’ needs to be pried open a little before even low-level production processes can be fully comprehended. This, it seems to me, parallels the emphasis in recent versions of Functional Grammar upon higher-level, top-down factors that co-determine –

in such a system is the maximal synchronized resonance of the representation of the component words within a syntactically conformal structure.
The non-linearity of speech production

along with the grammar code itself – the output of complex intentions to produce an utterance.

References


A speaker/hearer-based grammar
The case of possessives and compounds*

Theo A. J. M. Janssen
Vrije Universiteit, Amsterdam

This chapter argues for a speaker/hearer-based grammar conceived of as a structured set of procedures and procedural elements which clarify how interlocutors can produce and understand utterances. Some basic issues of a speaker/hearer-based grammar (primacy of the utterance, interaction between utterance and frame of reference, morphemes as procedural elements, the cueing function of phonetic elements in communicative processes) will be elucidated by an analysis of possessive constructions and novel compounds. Their use is based on a relationship that the speaker and hearer can integrate into their conceptualization of the frame of reference in the light of their interpretation of the current situation, their knowledge of the world and possibly the situations communicated in previous utterances and sometimes even utterances to come.

1. Linguistic assumptions

What is meaning and what is the locus of meaning? If utterances are meaningful, to what elements can we ascribe their meaningfulness? To their elements, called morphemes? What, then, is their meaning? If verbal utterances comprise morphemes, the meaning of utterances must be based at least partially on the meaning of morphemes.

* I would like to thank Ad Foolen, Mike Hannay, Wim Honselaar, Evelien Keizer, Bob Kirsner, Frederike van der Leek, Francisco Ruiz de Mendoza, Esther Pascual, Gerard Steen, Eva Tol-Verkuyl, Naoaki Wada, and the reviewers for their valuable questions, suggestions, and objections.

1. A morpheme is taken to be a lexical or a grammatical form/meaning pair (compare Bloomfield's notion of taxemes, [1933] 1969:166), including phrasal lexical items (Wray 2002), constructions, and meaningful patterns of intonation (e.g., Levelt 1999:108): “The intonational phrase (IP) is a sense unit”; see also Cutler and Hagoort (1999:150–151).
How do morpheme meanings contribute to utterance meanings? What is the division of labour between the two types of meaning? How do they differ?

The latter questions imply that both utterances and morphemes have meanings. But are phrasings such as “utterances or morphemes have a meaning” and “the meaning of utterances or morphemes” theoretically well-founded? What someone means by producing an utterance can only be determined by an interpreting human being, and even then just approximately, since “[a] person’s dispositions to accept sentences do not determine a unique interpretation of these sentences” (Rorty 1972:443). Therefore, when an expression such as meaning of the utterance or utterance meaning is used, it has to be read as an abbreviation of “what the utterer means or meant by producing the utterance involved” (i.e. the utterer’s meaning) or “what the addressee interprets as being the meaning of the utterance involved” (i.e. the addressee’s meaning). The same holds for the referentiality of utterances and morphemes. Verbal elements cannot by themselves refer. It is a human agent who is able to (intend to) refer to something by using verbal elements.2

A linguistic theory based on the practices of speakers and hearers cannot but adopt the primacy of utterer’s meaning (e.g. Grice 1968; Searle 1978; Suppes 1986) and should primordially be oriented to describing the production and understanding of utterances (Gardiner 1932; Levelt 1989:29). Although the phonological and semantic aspects of utterances are equally important in a speaker/hearer-based grammatical theory, this chapter will concentrate on the semantic aspects of utterances and particularly address the semantic aspects of the linguistic means involved.

When trying to understand what a speaker means, we try to determine how our interpretation of the utterance involved is compatible with our interpretation of the current frame of reference. We assume the speaker to mean what we think we ourselves would mean if we were in the speaker’s position and produced the utterance in the light of our interpretation of the current frame of reference. In usual speech situations, we understand utterances without consciously analysing them into their parts or meaningful aspects. As a rule we are even unable to indicate what part or aspect of a spoken utterance was understood first or later. The process of understanding an utterance is hardly based on analysis, taken as a conscious act. The same goes for the process of producing utterances. In ordinary spontaneous speech situations we usually even produce utterances without being aware of constructing them at all.

When producing an utterance, speakers generally assume that their addressees share their interpretation of the current frame of reference, at least largely. Nevertheless, hearers themselves have to interpret what qualifies as the current frame of reference to which the utterance is related.3 In other words, when a speaker says The

---


3. For the various aspects of a frame of reference, see Connolly’s (2006) classification of context.
A speaker/hearer-based grammar

ball is under the table – and not, for instance, The table is above the ball – the speaker interprets the current frame of reference in which the “objective” situation is viewed\(^4\) and the hearer has to interpret the presumable current frame of reference in which the relevant situation has to be conceived of.\(^5\)

When hearing an utterance, the addressee has to determine what, in all likelihood, was meant and why it was this that was meant. The phonetic structure of an utterance is just one factor which enables the hearer to interpret what the speaker means by the utterance, including its raison d'être. The hearer takes the phonetic structure as a cue to determine the morphemes that are to be considered parts of the utterance. The morphemes distilled as being possibly applicable are taken as hypothetical signposts\(^6\) in order to determine both what situation is meant by the speaker and why.\(^7\) The other factor is the current frame of reference to which the hearer assumes the utterance to be related. However, what counts as the current frame of reference is not pre-established, but has to be determined by the addressee.

Let us consider an example. A man carrying a child on his arms and jostling through a crowd says “My child needs help!” . Anyone who considers themselves as being addressed has to establish the utterance’s frame of reference on the basis of the non-verbal data available, and to match them with the phonetically shaped data of the utterance. Due to my, child will not be understood as a means of mentioning just a young person but as a means of mentioning a son or daughter of the speaker; needs will be interpreted as being used to indicate emergency, and help to indicate medical care. Considering what can be seen and heard, the addressees will take the speaker’s phonetic noise as being used to focus their attention (see, e.g., Levelt 1999:90) on his and his child’s situation, to announce his jostling, to request that the addressees let him through, to explain his jostling and also to excuse himself for it. The newly established frame of reference comprises the fact that the speaker has said “My child needs help!” , to be taken by the addressees as his statement that his child needs help, his announcement of his jostling, his request to let him pass, his explanation of and excuse for his jostling.

---

\(^4\) Levelt (1999:93) considers ‘perspective taking’ to be “at the very core of all conceptual preparation for speech”.


\(^6\) I use signpost to indicate an entity which is taken to serve as a morpheme in the phase at which the hearer has not yet determined how it is to be applied (or which of its so-called senses is applicable). A signpost itself does not say, prescribe, or instruct how to go; it is the person making use of it who, following an encompassing idea, decides which of all the directions assigned on it should be taken.

\(^7\) Compare how the relationship between utterances (messages) and morphemes (signs, i.e. signal/meaning pairs) are conceived of by sign-based theorists such as Diver (1969, 1995), Kirsner (1980, 2002), and Reid (1991).
Out of the context, parts of the phonetic stream represented by *my-child-needs-help* can serve in various ways, and very differently from the interpretation given above. *My child* could be used as a form of address (as in: "My child, [don’t be afraid!]"), where the addressee can be an adult and the referents of *my* and *child* do not need to be relatives (which is plausible, however, in the example). *Child* could be used as a part of the creative nominal compound (nonce word): *child needs*. *Needs* could be used as a noun, adverb or verb, and *help* as a noun or verb. Addressees may be assumed to be usually not aware of such possibilities, except when they cannot hear the phonetic stream well and/or cannot survey the situation involved adequately (and even then, they may be assumed to be aware of just a few of all the possible alternatives).

In the situation given, the hearers distil from the phonetic stream a morphemic structure which could be different in another situation. They can start the determining process before they have heard the phonetic stream of the utterance in its entirety. For instance, when hearing the phonetic stream *my child* and seeing the child on the man’s arms, they can already hypothesize (i) that the sound can be analysed as two morphemes, namely *my* and *child*, (ii) that these morphemes serve to refer to the speaker and the child on his arms and not to address the child or some individual in the crowd, and (iii) that the relationship indicated by means of *my* can be taken to be a kinship relationship. In other words, addressees normally do not first select a fully assembled morphemic structure (including just one semantic structure out of all the various semantic possibilities) and verify subsequently whether it is applicable to the situation concerned. The usual assemblage can be assumed to be based step by step on the hypothetical consecutive phonetic parts to be associated with morphemes (Cutler & Clifton 1999:142).

The linguistic and non-linguistic data, which have to be matched for an utterance to become understandable, imply a strong interdependence between the act of relating the phonetic structure to compatible morphemes and that of conceptualizing the frame of reference to which the realization of the phonetic structure is likely to be related (Langacker 1987:425–436; Hagoort et al. 2004).

How are we to conceive of the relationship between the meaning of an utterance and the current frame of reference, in view of the fact that the interlocutors by speaking and listening participate in and contribute to the incremental frame of reference? Both the situation meant and the reason why it is meant to be focused on are to be seen

---

8. The phonological and semantic determination of a phonetic stream such as represented by means of *John* and *Johns* in *Johnson* versus, for instance, *John’s ill* and *John’s illness* is assumed to be open during a fraction of time (e.g., Cutler & Clifton 1999:133–141). When processing *need* as a potential phonological-semantic unit (morpheme), hearers may be assumed to leave its semantics undetermined during a fraction of time. A similar short time of undeterminedness can be assumed in the case of processing *-s* in *needs*; the morpheme combination *need* + *s* can serve as a noun, adverb, or verb.
as inextricably interwoven with the frame of reference. Interlocutors mentally live in the frame of reference which they largely share and co-construct. That is why speakers do not portray the situation in question together with the communicative goal of their utterances, and hearers do not make a mental representation of what speakers intend to communicate. What happens is that speakers offer the phonetic cues which enable their addressees to determine what situation matters, and why. In Sinha’s (1999:232) words:

Acts of linguistic meaning […] are subjectively constructed so as to make sense in an intersubjectively shared universe of discourse, which is continuous with (not separate from) the material world in which other (non-discursive) human activities are carried out. This is the shared world of joint action, joint attention, and joint intention (Shotter 1993). This means that, rather than seeing acts of meaning as being about transferring mental representations from one individual to another, I see them as being about foregrounding matters of mutual concern in a given communicative situation.

The reason why a speaker draws attention to a situation is often hardly indicated, if at all. For instance, the illocutionary force of an utterance is usually not expressed explicitly, but is easily understood. The successful interpretation of the aim of an utterance is partly due to the fact that the speaker gives the necessary cues for determining the situation concerned, and partly due to the hearer’s knowledge about the circumstances in which an utterance of the relevant type is or can be produced (see Sinha 1999:234).

An utterance can thus be regarded as a non-accidental phonetic structure by means of which the speaker intends to guide the hearer in determining the situation meant within the current frame of reference, which – or rather, whose interpretation – the speaker assumes or can generally assume to be largely shared by the hearer. To some extent, the hearer has access to what is meant by the speaker (i) by perceiving the phonetic structure which is to be recognized as relatable to a string of morphemes serving

9. The idea that the process of understanding is oriented to its context, particularly, the utterer’s intention, has already been pointed out by Wegener ([1885] 1991:45; quoted by Nerlich & Clarke 1997:374): “And therefore to ”understand words” means as much as: to infer or guess their sense from the context. And one always needs to make some concessions for that: one must want to understand the remarks of an orator or author in his spirit”.


[meanings] do not add up “arithmetically” to the sum total of the message being communicated […] Rather, the meanings are no more than a collection of hints offered by the speaker, on the basis of which the hearer makes a guess at the message intended. […] The success of the communication thus depends to a large extent on the speaker’s ability to assess how much knowledge the hearer already has concerning the intended message, and what hints should be selected for a successful transmission of the new material.

See also, e.g., Kirsner (1980), Reid (1991:95), and Contini-Morava (1995:5–6).
Theo A. J. M. Janssen

as signposts to singling out the situation meant, and (ii) by taking into account, as best he or she can, what can be expected to be meant in the current frame of reference. Thus, the phonetic structure of an utterance does not in any form convey a mental picture or mental representation. Rather, it constitutes a perceptual basis that the hearer associates morphemes with, thereby taking into account what can be assumed to be the speaker’s current frame of reference. Understanding an utterance is an integrating act based on both bottom-up (from the phonetic structure upwards) and top-down (from the conceived current frame of reference downwards) informational procedures (Hagoort et al. 2004). The phonetic structure is a complex of cues by means of which the addressee is guided to understand the utterance (Gardiner 1932:69; Wittgenstein 1953:§85, §198; Reid 1991:40).

In the present view, the notion of sentence meaning is theoretically of no use. When determining a phonetic structure, the hearer has to render it into forms which qualify as the phonological parts of morphemes. The success of rendering the phonetic structure into a morphemic structure depends on the compatibility of the putative relevant morphemic structure with a situation fitting the current frame of reference. Hence, under this view there is no entity such that it can be considered to be the approximately (see Rorty 1972:443) definitive morphemic structure before its compatibility with a particular situation fitting in the current frame of reference has been achieved. It is not until the most plausible morphemic structure is established that the utterance meaning is available. In other words (see also Récanati 1998:625), there is no need for some intermediate interpretational level such as that of sentence meaning or “a (generally non-propositional) logical form”, as Carston (2002:64) assumes (following Sperber & Wilson 1986:175–176). As soon as one can trust the hypothetical

---

11. Langacker (1997:236) states: “a term such as ‘representation’ does not imply the existence of a picture-like image engraved in the brain; the term refers to “what the brain does, i.e. to recurrent patterns of neurological activity (or to aspects of the mental experience it constitutes)”. Reid (1991:7–8, 40, 314, 343–345) rejects the conception of language as a representational system in favour of language as a communicative instrument. Fowley (1997), who rejects mental-representation approaches to meaning, discusses some alternatives. For an overview of mental-representation approaches, see Nuyts (2001:Ch. 1).


13. The morphemic (lexicogrammatical) structure is constituted by the morphemes of an utterance, including the intonation into which the lexical and grammatical morphemes are integrated.

14. In Fauconnier’s (1994:2) words: “The speaker-listener does not consider all the interpretations of a sentence and then discard the inappropriate ones. He sets up a space of configurations starting from the configuration already available at that point in the discourse”.

interpretation of the entire utterance, one can trust the hypothetical interpretation of
the parts of that utterance, and vice versa.

The assumed morphemes of an utterance are taken to be signposts by means of
which the addressee is guided when determining the possibly relevant situation in
the current frame of reference. Morphemes do not function independently: "Out-
side a context a word has no signification at all (nor any denotation), but has [...] partly delimited possibilities or ranges of signification that are adapted, fitted, to context" (Ross 1981:37; Allwood 2003). In my view, the decision as to which morphemes
are used to constitute an utterance hinges on the way in which they are related to
the conceptualization of the current frame of reference, and more specifically to the
situation meant.

The semantic aspects of the morphemes which qualify as contributing to the
meaning of an utterance are not specific parts of some mental picture resulting from
the determination of an utterance meaning. The idea that a mental picture should dis-
play an utterance meaning is objectionable, because it requires an interpretation of the
mental picture and so on, which would lead the interpreter to an infinite regress of
reading mental pictures.15 What is more, the mental picture hypothesis distracts from
the fact that an utterance is an act.

Summarizing, I regard a verbal utterance as a cognitive-pragmatic phenomenon
by means of which the speaker attempts to enable the hearer to determine a partic-
ular situation (including the reason why attention should be drawn to it) within the
hearer’s conceptualized frame of reference. The hearer has to determine the situation
on the basis of a phonetic structure and a morphemic correlate that is compatible not
only with the phonetic structure but also with what the hearer conceptualizes as the
current frame of reference. In order to further clarify the cognitive-pragmatic status of
utterances, I will now proceed to sketch a speaker/hearer-based grammatical theory.

2. A speaker/hearer-based grammar

The primary objective of a speaker/hearer-based grammar is to characterize the pro-
cedures which interlocutors use in order to communicate. Such a grammar consists
of a set of procedures which enable interlocutors to produce and understand utter-
ances. The procedures do not serve to deliver mental representations of utterances or
to predict the possible sentences of a language.

A speaker/hearer-based grammar comprises four main procedures, namely two
complex acts performed by the speaker and two complex acts performed by the
hearer, see Figure 1. The two central speaker-based procedures are the specifying and
the vocalizing act; the two central hearer-based procedures are the perceiving and


- a “dual grounding” view of language and linguistic cognition, one which recognizes the centrality of both embodied grounding and functional or discursive grounding. As I see it, part of the underlying motivation for this view is a desire to escape from the restrictive scientific orthodoxy which says that you can have either an individual-cognitive, or a social-functional, view of language, but you can’t combine the two […]).
  (Sinha 1999:249)

These problems arise from the fact that objective theories of cognition are themselves the interpretations of analysts and hence are vacuous without an interpreter to give them meaning again. Terms such as ‘relevance’, ‘manner’, etc. are themselves interpretations of observed behaviour, and as such they need someone to make sense of the interpretation.
  (Robinson 1997:257)
principled necessity that speakers and hearers select the morphemes of their utterances in close connection with their interpretation of the current frame of reference (in which they participate). The question of how morphemic structures are assembled and analysed will be discussed in Sections 3 and 4 into some detail, where the use of possessive constructions and compounds are considered.

The speaker-related grammatical procedures of the specifying act are based on a structured set of morphemes, seen as two-part procedural categories (each category consists of a semantic and a phonological part), as elucidated with regard to My child needs help!. Considering the current frame of reference and the set of semantic categories, the speaker specifies what he or she means to communicate. The selection and integration of the suitable categories result in a morphemic structure. The phonological part of the morphemic structure is vocalized by the speaker into a phonetic structure, i.e. a string of sounds (or their written counterparts\(^\text{18}\)). The specifying act can still be assumed to be going on when the vocalizing act has already started (Levelt 1989:24, 199–200).

When indicating the situation meant, the speaker selects morphemes by whose phonetic realization the speaker expects to guide the hearer to the detection of the situation in the current frame of reference, assumed to be shared in large part by the interlocutors. Phonetic structures should not be taken as instructing,\(^\text{19}\) but rather as cueing the hearer to interpret them, i.e. to associate morphemes with them; the hypothetical morphemes can be seen as Wegweisers “signposts” (Wittgenstein 1953:§198).

When perceiving the phonetic structure, the hearer tries to determine how to understand what the speaker means by uttering the structure. Usually, the determining act is already operative when the vocalizing act is not yet completed. In Figure 1, the two main boxes on the right indicate that a morphemic structure is to be ascribed to the phonetic structure. While ascertaining which morphemic structure fits with the phonetic structure, the hearer starts from (i) what seems to be the appropriate conceptualization of the current frame of reference and (ii) his or her inventory of semantic and phonological categories. Both the interlocutors’ conceptualizations of the frame of reference and their category inventories usually coincide to a large extent.\(^\text{20}\)

\(^{18}\) Generally, speaker can be read as ‘writer’ (‘utterer’), hearer as ‘reader’ (‘interpreter’). Furthermore, vocalizing can be read as ‘writing’, and phonetic structure as ‘orthographic structure’. For psycholinguistic differences between hearing and reading, see, e.g., Cutler and Clifton (1999) and Perfetti (1999).


\(^{20}\) Langacker (1997:242–243) observes: “In a speech event […] each participant’s construal of the situation includes the fact that the other also apprehends it in a certain manner”. As for the category sets, Langacker (1987:62) states: “no two speakers control precisely the same set of units (e.g. the same set of lexical items and conventional expressions), even if their speech patterns
Although neither a frame of reference nor its conceptualization by interlocutors is a part of the grammar, the frame of reference is a prerequisite for each of all four procedures, because interlocutors apply the procedures in a close interplay with how they conceptualize the frame of reference. The frame of reference to be conceptualized by interlocutors comprises the relevant situation, in its conceptualized or conceptualizable embedding, which includes the ongoing discourse and the ongoing narration. In the course of their conversation, interlocutors co-construct their interactions (cf. for example Jacoby & Ochs 1995) and negotiate their conceptualizations of the current frame of reference, while attuning particularly the semantic categories involved to the situation they want to refer to (cf. for example Firth 1995). In producing and interpreting utterances, interlocutors conceptualize the incrementing frame of reference and take the assumed view of their communicative partners into account by observing "the highest norm of communication" (Bartsch 1984b:370–371):

Express yourself in such a way that what you say is recognizable and interpretable by your partner in communication, to the highest degree in agreement with what you intend him to understand. Correspondingly, for the hearer: Interpret such that the interpretation will be to the highest degree in agreement with what the speaker intends.

I do not subscribe to the instructional implications of Bartsch's norm, since observing it is generally not dependent on the interlocutors' conscious considerations, but is rather an irrepressible act of those who use the language involved. I do adopt, however, her main idea that the hearer focuses on what can be recognized interpretively as to be the speaker's intention in the given circumstances. Interlocutors constantly re-conceptualize their frame of reference on the basis of their involvement in the ongoing circumstances. Situational data, including data based on one or more previous utterances, and data based on one or more previous morphemes in the same utterance, enable the hearer largely to anticipate the following information to be provided appear to be identical". See also Wotjak (1974:39–41), Uhlenbeck (1981:16), Dik (1997,1:10–11, 1997,2:410–412), Barsalou (1992), and Barsalou et al. (1993).

21. The interlocutors are assumed to have similar imaginative powers and ways of inferential reasoning, including those based on the Cooperative Principle (Grice 1975, 1978, 1981) and the Politeness Strategies (Goffman 1955; Brown & Levinson 1978).


by the speaker. What is to be conceived of as given is the basis for updating the frame of reference by means of the new information. Each piece of information may be assumed to facilitate the understanding of what comes next (Reichling 1935:410) and to contribute to the ongoing conceptualization of the relevant frame of reference.

The current stage of the frame of reference provides the hearer with an informative basis for determining how the different data serve to conceptualize the situation meant. And, conversely, both the verbal and situational (including paralinguistic) data are conceptualized as contributing to the incrementing frame of reference. Both the linguistic and non-linguistic data are assumed to serve as cues. When the linguistic data are interpreted as hypothetical morphemes, they serve as signposts by means of which the incrementing frame of reference is restructured. Neither the linguistic nor the situational data can a priori be considered to be interpretively clear-cut and ready-made. They are open to interpretation within the confines of the frame of reference at that point in the discourse. The frame of reference serves as a complex constraint (Cicourel 1992:307) on what can be said next and on how it can be interpreted. In conceptualizing the frame of reference one cannot take into account an infinite number of situational data, since interlocutors cannot know everything about the context.

Delimiting is an agent-based act.

The present view is concisely characterized by Gardiner’s (1932:34) dictum that “the function of words is to make the listener ‘see what is meant’. They are in fact ‘clues’”. In producing an utterance, the speaker uses morphemes whose phonetic realizations serve as cues for the hearer to determine what particular situation matters and why it matters in the frame of reference, whose conceptualization the speaker assumes to be largely shared by the hearer (see also Levelt 1999:90). Cueing the hearer occurs as follows. The speaker gives shape to the cues by using a phonetic structure based on a morphemic structure. The phonetic structure enables interlocutors to make perceptual contact. The hearer has to determine what morphemes may be assumed to be based on the phonetic structure, and has to figure out what morphemic structure is compatible with both the phonetic structure and what he or she assumes to be the speaker’s conceptualization of the frame of reference. It is obvious that, in this view,

24. Following Clark (1983:333, 335): “In interpreting complete utterances, listeners ordinarily infer a hierarchy of goals they believe the speaker is trying to attain, and they interpret the speaker’s current utterance as a step in the plan for attaining one or more of these goals”. And: “Listeners generally cannot, nor are they expected to, infer the speaker’s hierarchy of goals accurately without consulting both the utterance and their common ground”. See also Clark (1983:336–337). For a survey of discussions on the interaction between top-down and bottom-up information, see Cutler and Clifton (1999).

25. See also Hagoort et al. (2004) and Langacker (1997:237): “our understanding of a sentence does not occur all at once but evolves and changes during the course of its occurrence”. In Clark’s (1983:340) words: “Parsers must worry about the speaker’s intentions at every turn”.

the phonetic structure does not convey a mental picture or a thought, as Gardiner (1932:69) rightly notes:27

The sounds of speech are not aeroplanes invented for the purpose of carrying thoughts as their passengers between man and man […] The impossibility of transferring thought is absolute and insurmountable. Only by inference from his own thought can the listener conclude that the speaker has been thinking of the same thing. What passes in speech between the two persons concerned is mere sound, bereft of all sense.

If the phonetic structure cannot convey a mental representation of a situation, one wonders why a morphemic structure should be held to be some kind of an embryonic description of the situation meant. What might be the communicative function of a mental picture of how the speaker conceptualizes the relevant situation? And why should the hearer make a mental picture of what the speaker may be assumed to indicate in producing the utterance, particularly when the situation meant is interwoven with what can be taken to be the directly perceptible frame of reference?

Let us consider the act the hearer is supposed to perform in order to understand an utterance. In determining the phonetic information connected with what can be conceptualized as the current frame of reference, all the hearer needs is a few indications so as to be able to see what situation the speaker means to refer to and why.28 On such a view, “linguistic communication is more a matter of finding one’s way in a context than of exchanging messages in some abstract code which is determined a priori and independently of contexts” (Bosch 1985:143). Thus, no mental picture, conceptual, or semantic representation is assumed to be conveyed or portrayed by means of an utterance. The morphemic structure serves for the speaker to specify – for the hearer to be taken as specifying – the signposts guiding to the situation in the incrementing frame of reference.

The interconnection between the choice of morphemes and the frame of reference has been studied by Olson (1970), who proposes “a cognitive theory of semantics” which “necessarily involves the language user’s non-linguistic knowledge by showing that the choice of words in an utterance is a function of […] the speaker’s knowledge of referents” (Olson 1970:258). Words are used to “specify perceived events relative to a set of alternatives” (Olson 1970:263). He elucidates his view by means of the following test. In three different situations, a subject saw that a gold star was placed under one and the same small wooden block. Then, the subject was instructed to say to another person, who did not see the act, where the gold star was placed. The situations and answers are displayed in Figure 2.

27. Compare Reddy (1979) and, e.g., Reid (1991:7–8).

28. See Fillmore’s (1977b:86) dictum: “A word or phrase or sentence or text identifies a scene, and it foregrounds, or highlights, some portion of it”.
A speaker/hearer-based grammar

The test showed that in every situation the subjects specified the relevant object to the level required by the hearer to distinguish it from the alternatives in the current frame of reference. Olson (1970:266) concludes that "the level to which an utterance is differentiated depends on the intent of the speaker. No utterance ever differentiates an intended referent from all possible alternatives but only from those among which the speaker infers that the listener must choose in the present context, for the purpose of that particular 'language game'". This pragmatic view of the differentiating depth of utterances has an interpretive counterpart for the actual hearer (in conformity with Grice's maxim of quantity) and, as Gardiner (1932:52) suggests, for the theorist: "it seems necessary to regard the things meant by speech as [...] susceptible of never-ending analysis, but we must add as a rider that the theorist of speech is only concerned with so much of those things as is required to elucidate what the speaker intended the listener to see".

Olson's cognitive theory of semantics assumes the meaning of an utterance to be the information serving as a linguistic signpost which reduces the set of alternatives to the event meant: "Since not all the potential information of the referent as perceived is signalled by a word, we conclude that words (or utterances) neither symbolize, stand for, nor represent referents, objects, or events. They serve rather to differentiate some perceived event from some set of alternatives" (Olson 1970:265). One might ask oneself how one can deal with situations which are not directly perceived by one or both of the interlocutors. Olson (1970:268) suggests that "the primary extension of the ac-

---

29. Olson (1970:266) is aware of the fact that "speakers frequently call a cat a 'cat', even if for informative purposes 'animal' would be sufficient". See Rosch (1973) and, e.g., Levelt (1989:129–134, 223) for the preference of 'basic level terms'. See Fillmore (1985:234–239) for the question of "Why did the speaker select this form in this context?". See Levinson (2000:33) for the heuristic "What is said in an abnormal way, isn't normal; or marked message indicates marked situation". Francisco José Ruiz de Mendoza (p.c.) suggests adding the following heuristic: 'What is said is a function of what is distinct in context'. See also Ariel (1988) and Robinson (1997:257–268).

30. Cicourel (1992:309) discusses the "nagging issue" of an infinite regress when an observer of speech tries to determine the relevant context. See also, e.g., Nunberg (1978:40–42) and Robinson (1997:258–259).
count to these events would be to hypothesize that the speaker makes his semantic decisions on the basis of inferred alternatives. The speaker is assumed to “make an estimate of these alternatives both in the light of the preceding utterances in the discourse and from the experience of the listener” (Olson 1970:266), although there is “no time to consciously weigh the alternatives before deciding on a word” (Levelt 1989:22). From the point of view of the determining procedure, the hearer may be expected to be able to contextualize the cues in such a way that the frame of reference needed is inferred and adapted adequately.

I have argued that the notion of “sentence meaning” is of no use from a theoretic point of view. When determining the phonetic structure, the hearer has to convert it into forms which qualify as the phonological parts of morphemes. The success of converting the phonetic structure into a morphemic structure depends on the compatibility of the morphemic structure that is taken to be relevant with a situation fitting in the possible current frame of reference. There is no entity that can be considered to be the definite morphemic structure (including the definite semantic structure) before its compatibility with a particular situation fitting in with the current frame of reference is established. It is not until the (approximately) definite morphemic structure is established that the (approximately) definite hearer meaning is available. Thus, no intermediate sentence meaning is established apart from as well as different from the utterance meaning.

In order to elucidate the relevance of my approach to utterances, which is based on both interlocutors’ verbal interaction and their mutual frame of reference, I will discuss two types of contextual expression, namely possessive constructions and novel compounds, since “[w]ith contextual expressions, reference to the speaker’s and hearer’s common ground is mandatory” (Clark 1983) 1992:337. The use of such elements cannot be processed by an automatic procedure and cannot be formalized by means of a blindly operating algorithm (see for example Leech 1983:36; Contini-Morava 1995:6; Seuren 1985:19–21, 314). Although possessive constructions and novel compounds might be considered paragons of “contextual dependence”, I assume that most, if not all morphemes require that the hearer attunes them similarly to the frame of reference involved.


32. See the parametrization approach to for instance good, as proposed by Ruiz de Mendoza Ibáñez and Santibañez Sáenz (2003) and Ruiz de Mendoza Ibáñez (forthc.). In this view, speakers and hearers consider an entity to be good if it is positively assessed by them. Whether their assessment occurs within in the scope of, e.g., morality, pleasure, or admiration is a matter of parametrizing their use of good in accordance with textual and contextual clues. Good in She is a good nun could be interpreted as ‘virtuous’, ‘efficient in doing her job’, or ‘kind’. The
3. The case of possessives

The genitive, as in John’s dog, can be taken to be monosemous,\(^{33}\) although John’s dog can be used to refer to the dog John possesses, or the dog John is standing in front of, the dog John saw yesterday, the dog John always wanted, and “any number of other things” (Clark 1992:311). Interpreting possessives “relies heavily on the coordination of the speaker and addressees. Possessives, in short, are contextual expressions” (Clark 1992:311).

The interpretive variety of genitival expressions – and other possessive expressions such as his dog and the dog of that man\(^ {34}\) – can be based on the semantics of \(x’s\ y\), namely “\(y\)” is related to “\(x\)”, whereas how “\(y\)” is related to “\(x\)” must be inferred from the current frame of reference. This analysis (Janssen 1975, 2003) originates from Bendix’s (1966:57) observation that possessives can be used when very casual relationships are involved, such as in (1).

(1) A bull chased me and a bull chased John. My bull was fiercer than John’s bull.

In the first utterance in (1), the bulls and the persons are asserted to be in a chaser-chasee relationship. In the second utterance, the chaser-chasee relationship is not expressed. Though it is indicated by means of the possessives that the bulls are related to two persons, we do not know how they are related independently from the information about the event referred to before. We may assume that the chaser-chasee relationship cannot qualify as a sense of a possessive pronoun or a genitive, since it is an entirely accidental relationship.

In order to determine the semantics of the possessives in (1), let us compare (1) with (2).

(2) A bull chased me. The bull was fierce.

By using the bull, the speaker refers to “the bull which chased me”. The function of the possessives in (1) resembles the function of the definite article in (2). Both the possessives and the definite article are used to indicate a reference relationship.\(^ {35}\) When


34. The relationship between prenominal and postnominal possessives in English are discussed by, e.g., Taylor (1996), Willemsen (2003), and Keizer (this volume).

35. Some remarks. So-called possessive structures do not always serve to express a reference-point relationship. In, e.g., My dear, dear little child!, it is not due to the use of my that the child
using such elements, the speaker needs a frame of reference within which the entity in question is unique (Janssen 1975). As to the second utterance in (1), the speaker conceptualizes two sub-frames of reference, each comprising one bull and some person. When using *my bull* and *John’s bull*, the speaker specifies that the bull in one frame has to be seen as related to the speaker as being unique in it, and the bull in the other frame has to be seen as related to John as being unique in it (Janssen 1995).

The present view holds that the different possessive morphemes under consideration can be analysed as monosemous. Why, then, is a combination of possessives such as in (3) and (4) odd?

(3) I see his car, son and nose.

(4) I see John’s car, son and nose.

The oddity of (3) and (4) seems to suggest that there are different possessive relationships involved in *his car, his son, and his nose*, due to the polysemy of *his*. Nevertheless, rather diverse possessive relationships are reconcilable. Consider (5) and (6).

(5) John likes to talk about his family, work, vacations, way of life, irritations, recent decisions, ideas, plans, and prospects.

(6) Can you tell me who John’s colleagues, neighbours, friends, teammates, and closest relatives are?

In (5) and (6), the entities indicated by means of possessive structures combine very well (Janssen 2003:110–114). In (5), the entities constitute a purely coincidental class of favourite subjects; in (6), the individuals in question are members of different social
networks, featuring a professional, demographic, affectional, sports, and family relationship with John. In both (5) and (6), the entities concerned share a more or less casual property which allows us to join them under one category. The coincidental property is so accidental that the words involved should not be assumed to have it as an aspect of their meaning. Apparently, it is the way we conceptualize the entities in the current frame of reference that allows us to categorize them as comparable.

How can one do justice to the intuition that an “extraordinary variety of relationships” (Langacker 1993:7) is covered by possessives? This issue relates to the question of how to select one out of all the possible relationships in an actual case. Caramazza and Grober (1976:199) reject the idea that language users have to search “through a list of senses until the appropriate one is found” (see also Diver 1969:46). In their view, the relevant value is not determined by finding a stored representation, but computed on the basis of a “core meaning”. In my view, the appropriate value is obtained on the basis of a morpheme’s semantics, serving as the initial state of the hearer’s interpretive procedure. In this procedure the hearer tries to match the semantics of the hypothetical morpheme and the hypothetical conceptualization of the current frame of reference.

In the first utterance in (1), A bull chased me and a bull chased John, the speaker asserts that there is a chaser-chasee relationship between two bulls and two persons. When producing the second utterance in (1), My bull was fiercer than John’s bull, the speaker relates this utterance to what is asserted in the previous one. How is the chaser-chasee relationship in the first utterance integrated into the incrementing frame of reference in which the situation of the second utterance has to be viewed? Consider Figure 3, which presents an encompassing frame of reference shared by the speaker and hearer, and two sub-frames: in one of them a bull is seen as uniquely occupying the speaker’s domain; in the other one a bull is seen as solely occupying John’s domain.

Let us focus on the conceptualization of the two sub-frames of reference applicable to the utterance My bull was fiercer than John’s bull. In order to produce this utterance, the speaker selects the possessive pronoun my and the genitive John’s, by means of which the bulls can be singled out, due to the fact that each of them can be construed as being unique in one of the two sub-frames of reference. In neither case is the accidental chaser-chasee relationship expressed by means of the possessives. By using the possessive morphemes, the speaker enables the hearer to relate one bull to the speaker and the other to John. It is due to the frame of reference available when My bull was fiercer than John’s bull is uttered that the hearer can infer a chaser-chasee relationship.

38. The adaptation of a morpheme’s semantics to the current frame of reference is taken as a parametrization procedure by Ruiz de Mendoza Ibáñez and Santibáñez Sáenz (2003) and Ruiz de Mendoza Ibáñez (forthc.).

39. The speaker could have said: The bull that chased me was fiercer than the bull that chased John, but a possessive construction is a more economic and compact manner to differentiate between the bulls (Weydt 1972:57). Relative and possessive constructions may be assumed to take different positions in the referential accessibility hierarchy (Ariel 1988).
Theo A. J. M. Janssen

in the case of my bull and in the case of John’s bull. This inference enables the hearer to conceptualize the two utterances in (1) as coherent.

In the present view, the chaser-chasee relationship asserted in A bull chased me and a bull chased John is reconceptualized as a relationship of “A in B’s domain”, expressed by means of the possessives in the second utterance. There are two arguments supporting the idea that the sub-frames of reference are not automatically produced on the basis of the chaser-chasee relationships expressed in the first utterance, but result from their reconceptualization.

First, the relationships which would have to induce the relationships indicated by means of possessive expressions are grammatically too diverse. Compare, for instance, (1) and (7).

(1) A bull chased me and a bull chased John. My bull was fiercer than John’s bull.

(7) I caught a fish and John caught one. My fish was bigger than John’s fish.

The subject and direct object referents of the first utterances in (1) and (7) do not relate, in the same order, to the possessors and the relevant possessees in the last utterances in (1) and (7). The way in which the chaser-chasee relationship corresponds to the possessor-possessee relationship in the last utterance in (1) is the reverse of the way the catcher-catchee relationship corresponds to the possessor-possessee relationship in the last utterance in (7). Therefore, it is implausible that the relationships between the entities referred to in the first utterances blindly give rise to the relationship to be expressed by means of possessives.

I assume that possessives are used to express a relationship between two entities in such a way that one of them is unique in its relationship with the other entity.

40. When, e.g., the son of the king is used, the unique reference to the son – interestingly, with family members and body parts, the reference need not be unique (pers. comm., Evelien Keizer) – is not due to the possessive structure, but to the presence of the definite article, as
It is the fact that an entity can be assumed to be the only one featuring this relationship with another in the current frame of reference which allows the first entity to be singled out. The relationship which is expressed by means of a possessive does not coincide with the relationship between the entities, which is cognizable on the basis of the frame of reference, particularly one or more previous utterances. Due to the fact that this relationship can be known very easily, the impression that the possessive is used to indicate this cognizable relationship thrusts itself upon the language user, but, again, the relationship that is expressed by means of a possessive is not identical to this cognizable relationship.

The second argument in support of reconceptualization is that a sequence of utterances such as in (8) shows that the first utterance does not automatically need to produce two sub-frames of reference which enable the speaker to refer to the two bulls in the last utterance.\textsuperscript{41}

(8) A bull chased me and a bull chased John. Both bulls were terribly fierce.

Suppose the first utterance in (8) evokes the conceptualization of a sub-frame of reference in which the bulls appear together as a unique set. Would, then, the first utterance in (8) also evoke the conceptualization of a set of two sub-frames of reference? The coherence of the first and second utterances does not need such a conceptualization. The coherence of the utterances in (1) or (7), however, does not require a conceptualization such as in (8). If such conceptualizations were part of the normal procedure, regardless of their functionality, they would lead to information overload.

Furthermore, let us consider (9), produced without earlier discussion of the current or a previous lecture.

(9) Yesterday’s lecture was less boring.

Should we assume that a sub-frame of reference in which a lecture is conceived of as a unique element relative to yesterday is already available before (9) is uttered? A positive answer is implausible, since an innumerable number of entities, including a great number of lectures, can be considered to relate to yesterday but the genitive does not automatically induce matching sub-frames of reference in advance. It is not until the act of specifying (and determining) yesterday’s lecture is actually being performed that the language user conceptualizes a frame of reference in which the lecture is related...
to yesterday as a unique element in this relationship. A great many lectures may have been delivered yesterday, but the lecture referred to is considered by the speaker to be the noteworthy one, for instance, because it was the only one both the speaker and the addressee attended yesterday. In order to determine which lecture is at issue, the hearer has to conceptualize a frame of reference to which the entity can be related as unique, in accordance with what is presumably the view of the speaker. This resembles what a viewer has to do when encountering the sign *Beware of the dog!* at the gate of a house. In cases like this, Hawkins (1984:656) points out: “finding the pragmatic [...] frame is all that the hearer can do, everything else he must take on trust from the speaker”. Knowledge of the world tells the viewer that he or she has to watch out for the dog related to the house with the sign outside it. Nobody, I imagine, would assume that the article *the* has a sense expressing this relationship.

In the present analysis of possessives, speakers and hearers are not assumed to search through a list of senses of possessive morphemes in order to decide which one is applicable in the relevant situation. They are assumed to conceptualize a frame of reference on the basis of the information that appears to be relevant at the current stage of the ongoing discourse. For instance, in the second utterance in example (1), the information about the chaser-chasee relationship, which is cognizable on the basis of the first utterance, is reconceptualized in view of the specifying (or determining) act being performed. The speaker selects the morphemes which are best suited to express the possessive relationships, without specifying what is already cognizable. Morphemes such as possessive pronouns and the genitive qualify for this task.

Although possessives are interpretively compatible with various cognizable relationships, they are not hyperonyms of the morphemes by means of which such relationships are specified in a previous utterance. The relationship that is compatible with a possessive morpheme must match what can be retroactively integrated into the conceptualization of the frame of reference being updated on the basis of one or more previous utterances or on the basis of the context of situation and knowledge of the world. It is the conceptualization of the current frame of reference that makes it superfluous for the hearer to seek the adequate interpretation by searching through a list of senses of a morpheme. Which interpretation is applicable is computed during the process in which the update of the incrementing frame of reference is conceptualized in relation to what is being specified in the utterance under construction and in relation to the cognizable information which is easy to obtain from elsewhere, i.e. world knowledge or knowledge based on a previous utterance.

In view of the analysis in which the possessive relationship differs semantically from the cognizable relationships with which it is compatible, it is obvious that the updating of the frame of reference related to the utterance planned to comprise a pos-

---

42. In Clark’s (1992:339) words: “Intentional parsers create senses and don’t just select them from a predetermined list of senses”.

sessive relationship cannot be processed by an automatic procedure. The processing required is agent-based, i.e. it has to be performed by conceptualizing human beings.

4. Novel compounds

In his discussion of contextual expressions, Clark (1983 [1992:311]) claims that “innovative compound nouns are contextual expressions” since their interpretations are not denumerable and depend on “the close coordination of the speaker and addressee”. His idea is in line with Jespersen’s characterization (1942:137); “Compounds express a relation between two objects or notions, but say nothing of the way in which the relation is to be understood. That must be inferred from the context or otherwise”. This or some similar idea is held by a number of linguists, such as Downing (1977), Bauer (1978), Kay and Zimmer ([1976] 1990), Meyer (1993), Ryder (1994:83), and Anstey (this volume).43

Clark’s and Jespersen’s characterizations of compounds have two points in common with the description of possessives presented in the previous section. Possessives and compounds share the basically open-ended interpretability and the interpretive dependence on the speaker and hearer, as observed by for instance Kay and Zimmer ([1976] 1990), Clark ([1983] 1992), and Meyer (1993:7–9). These two points will be the main themes of this section.

English compounds consist of two consecutive elements: the first is the modifier of the second element, i.e. of the head of the compound. The modifier narrows the semantic range of the head to a semantic subclass in such a way that the applicability of the head as a class term depends on its relationship with the modifier. The way in which the head is related to the modifier is not expressed, but is supposed to be known or to be inferable by the hearer. In the case of Downing’s (1977:818) well-known novel (so-called deictic) compound,44 used when a friend was asked to sit in the apple-juice seat, the class of seats was restricted to the seat in front of which a glass of apple juice had been placed. In the situation at hand, the speaker was not so much in need of a new class term for a name-worthy category as that she wanted “a demonstrative device, having more in common with a proper name than a class name” (Downing 1977:823).

43. See also Anstey (this volume) for a survey of the (psycho)linguistic literature on novel noun-noun compounds.

44. The difference between the need for a temporary situation-related referential device and the need for a new classificatory term can be taken to distinguish deictic compounds from other novel compounds (Downing 1977:939).
Here, I will examine compounds comprising a sentential modifier consisting of one or more fictive utterances (henceforth “utterances”). I will argue that such complex modifiers can be considered to present a fictive utterance or text rather than an entity without actual referential possibilities. Some compounds with a sentential modifier are, however, fully conventionalized, such as the following Dutch dictionary entries: *blijf-van-mijn-lijfhuis* ['stay-off-my-body house'] “women’s refuge centre” (Van Dale 1984), *doe-het-zelfwinkel, doe-het-zelfzaak*, both “do-it-yourself shop”, and *doe-het-zelfartikel* “do-it-yourself article” (Van Dale 1992), *lach-of-ik-schietshow* ['laugh or I shoot show'] “show full of broad humour” (Van Dale 1999).

Now, let us consider some characteristics of novel compounds with a sentential modifier. The modifier can occur “in all four established illocutionary structures: (i) declarative, (ii) interrogative, (iii) imperative, and (iv) exclamative”, as shown by Pascual (2002:206) with examples such as the following (Pascual 2002:208–211; my italics).

10. Declarative
   The problem is that we all rely on the “I’m the only one anyway” syndrome.

11. Interrogative
   I know why I wasn’t saying it, because then we would get into these do we understand each other? conversations.

12. Imperative
   Is he like these […] give me a shot at everyone in the world at once people?

13. Exclamative
   so we just had this oh, God this is the worst party we’ve been at for years feeling.

Sentential modifiers can comprise coordinated clauses or a main and a subordinated clause. Consider the following Dutch cases (Pascual & Janssen 2004; abbreviations: AV (Ad Valvas), NH (NRC Handelsblad), VG (VPRO Gids); my italics).

   ‘Driven by the humiliation and the feeling of insecurity, he zigzags between the missionary imperialism of the neoconservatives and the “warn-hit-and-leave” nationalism of the classical conservative Republicans.’

---

45. *Sentential* regards here the presentation of one utterance or a combination of utterances. See Pascual and Janssen (2004) for an overview of studies on compounds with a sentential modifier.

46. For the linguistic notion of fictivity, see Talmy (2000,1:100–103), and particularly Pascual (2002).
(15) Het verkeer raasde rond met een *pak-me-dan-als-je-kan*-geestdrift. (Basha Faber, *De jeugdzonde*, Amsterdam: Augustus, 2002, 135)
'The traffic raced round with a catch-me-if-you-dare passion.'

(16) Marginale *wie-niet-sterk-is-moet-slim-zijn*-komedie van David S. Ward (King Ralph) met Kelsey Grammer als inventieve duikbootkapitein Dodge. (VG, 25/11/03)
'A marginal *he-who-is-not-strong-must-be-clever* comedy by David S. Ward (King Ralph) with Kelsey Grammer as the inventive submarine captain Dodge.'

In (14) the modifier consists of three coordinated clauses; in (15) it consists of a main clause combined with a conditional subordinate clause; in (16) the subject of the main clause is a headless relative.

Furthermore, sentential modifiers can also comprise two or more asyndetically connected clauses constituting a brief text (piece of discourse). See, for instance, the cases in (17)–(22) (apart from the latter, the examples are from Pascual & Janssen 2004; my italics).47

(17) *Want het moeilijk-saai-ik-kan-er-niks-meebeeld van jongeren is niet hun schuld.* (NH, 22/10/03)
'After all, the *difficult-boring-that's-no-use-to-me* picture of youngsters is not their fault.'

(18) *Met de "Niet tevreden? Geld terug!"-garantie kunt u zonder risico het Braun FreeGlider of Syncro scheerapparaat uitproberen.* (Braun’s information leaflet, 2003)
'Due to the *Not satisfied? Money back!* guarantee, you can try the Braun FreeGlider or Syncro shaver without any risk.'

'*Drive carefully – Hands on the wheel* action: Eileen gives away driving trainings to young automobilists!'

(20) Wetenschap moet niet proberen aan te sluiten bij publieke stereotypen met een hoog *zie-je-wel-dat-vond-ik-altijd-al-gelijke* […] (NH, 27/09/03)
'Science should not try to tie up with public stereotypes featuring a high *you-see?-that-is-what-I-always-thought* content.'

(21) *Ze getuigt […] van een "u vraagt, wij draaien"-mentaliteit.* (AV, 12/02/04)
'She shows a "you ask, we play" mentality.'

---

47. The compounds in (17) and (18) are deictic in Downing’s sense, but the compounds in (20)–(22) can be taken to be class terms, due to the use of the indefinite article *een ‘a’ and a kind of*. The example in (19) is a case in-between.
(22) I have never been in a situation even remotely like this before, and I do have a kind of “who needs THIS? – lemme outa here!” reaction. (E-mail, 04/06/04)

The modifier in (17) consists of three declaratives (the holophrastic “utterances” moeilijk and saai, followed by the “utterance” ik kan er niks mee). The modifier in (18) shows an interrogative and a declarative elliptical “utterance”, the latter serving as a commissive. The modifier in (19) comprises two imperatives, the form of the latter being elliptical. In (20), an interrogative is combined with a declarative. In (21), two declaratives are combined. And in (22), an interrogative is combined with an imperative.

The examples in (10)–(22) show (i) that compounds with a sentential modifier are a highly elaborated and fully productive category, both in English and Dutch, and (ii) that the sentential modifiers serve as “utterances”, i.e. fictive intentional entities.

The main point of the intentionality of the modifiers is that conceptualizing human beings are involved in producing and understanding them in order to attain a certain goal. Compounds with a sentential modifier need at least a twofold conceptualization: what has to be conceptualized is the modifying “utterance” (or “text”) and the way in which the modifier is to be related to the head.

Let us consider two examples as a means of illustrating that a twofold conceptualization is required for producing and understanding compounds with a sentential modifier. The following Dutch example is taken from Sesamstraat (“Sesame Street”).

(23) Dit is niet de honkbalspeelklop maar dit is het jij-bent-mijn-goeie-vriend-en-ik-de-jouwe-klopje. (NPS-Nederland 3, 05/01/04; my italics)

‘This is not the baseball knock but this is the you’re my good friend and I’m yours knock.’

What is the context of situation? Telly knocks at Oscar’s “door” to play baseball with him. This leads Oscar to define a single knock as a hoepelklop “buzz off knock” and a double knock as a honkbalspeelklopje “baseball game knock”. Teased by these tricky “conventions”, Telly wonders if Oscar is still his friend. When again he knocks once, he outsmarts Oscar by immediately calling his single knock a jij-bent-mijn-goeie-vriend-en-ik-de-jouwe-klopje.

We can observe two important points. First, the modifier jij-bent-mijn-goeie-vriend-en-ik-de-jouwe is an “utterance” in which the pronouns jij “you”, jouwe “yours”, ik “I”, and mijn “my” are used exophorically to refer to the interlocutors Telly and Oscar. Telly even points at himself when pronouncing ik as a part of the compound. This implies that the pronouns serve as indexical elements across the compound’s bound-

48. Langacker (1999b: 90–91) takes the utterance Who needs that car? as a case of a virtual speech act in which “the speaker only pretends to ask a question. The actual interactive intent is not to elicit an answer from the hearer, but to render evident the impossibility of providing a truthful answer that satifies the question’s existential presupposition”.

49. For Afrikaans, see, e.g., R. Botha (1981); for German, see, e.g., Meibauer (2003).
The second point is our understanding of the relationship between the modifier and the head. In the previous context of situation, Oscar defined the communicative functions of two types of knock. However, in the current frame of reference it is Telly who defines his single knock by means of the compound \textit{jij-bent-mijn-goeie-vriend-en-ik-de-jouwe-klopje} as a knock in order to communicate: you are my good friend and I am yours. Due to the fact that the word \textit{klopje} “knock” can be used to refer to a noise produced to draw someone’s attention and even to signal something, it facilitates the idea that the relationship between the head and the modifier is that the knock serves to express “you are my good friend and I am yours.”\footnote{As Bob Kirsner pointed out to me, the diminutive \textit{klopje} might have a cue-value, aiding the communication, since it can suggest intimacy (non-threateningness), whereas the diminutive \textit{hoepelopklopje} ‘little buzz off knock’ sounds weird.} Therefore, both linguistic and situational data can be assumed to contribute to a successful verbal interaction between speaker and hearer. The processing required has to be performed by conceptualizing human beings.

Let us now consider the following Afrikaans example, in which the complex compound contains an English sentential modifier.

\begin{quote}
\textit{Mbeki se I am an African-toespraak is ‘n goeie illustrasie hiervan. (Willem Botha, Identiteit en taal. Aambeeld / Anvil 30-2, 2002, 13)}\footnote{The metaphorical implications of Thabo Mbeki’s statement “I am an African” are explained by W. J. Botha (2001).}
\textit{‘Mbeki’s I am an African-speech is a good illustration of this.’}
\end{quote}

We will consider three important points. First, the modifier in \textit{I am an African-toespraak} is an “utterance” in which the pronoun \textit{I} is used to refer to Mbeki (Deputy President of South Africa when he gave an address at the time of the adoption of South Africa’s 1996 constitution bill), as is explained in (25), immediately following (24):

\begin{quote}
\textit{In die genoemde toespraak vereenelwig Mbeki hom deur die gebruik van die eerstelpersonvoornaamwoord \textit{ek} baie intiem met verskillende subidentiteitie van ‘n multikulturele en multi-etniese Suid-Afrikaanse samelewings deur onder andere die volgende uitdrukings te gebruik: “I owe my being to the Khoi and San […] I am formed of the migrant who left Europe […] I am the grandchild of the warrior men and women […] I am the grandchild who lays fresh flowers on the Boer graves of St Helena […] I come of those who}
\end{quote}

\footnote{By using \textit{THIS} in (22), the writer refers to the predicament she has mentioned in the first conjunct; by using \textit{me} the writer refers to herself; by using \textit{here}, she refers to the situation indicated in the first conjunct.}
were transported from India and China [...] (Willem Botha, Identiteit en taal. Aumbeeld / Anvil 30-2, 2002, 13)

‘In the speech mentioned, Mbeki identifies himself by means of the first person personal pronoun *I* very closely with various subidentities of a multicultural and multiethnic South African society by using expressions such as the following: […]’

The passage in (25) illustrates that the pronoun *I* does serve as an indexical element across the compound’s boundary. Again, we may conclude that a compound does not encapsulate the indexical potential of a pronoun when it is a part of a modifier (*pace* Bresnan & Mchombo 1995:193–194).

What’s more — and this is the second point — the fact that different languages are involved does not impede coreference. The element *Mbeki se* is Afrikaans, while *I* is part of the English modifier. Certainly, it is implausible to import English syntactic rules into Afrikaans (word formation) rules in order to account for examples such as *I am an African-toespraak* (Bresnan & Mchombo 1995:194). However, there is no need to build English grammar into the grammar of Afrikaans. Bilingual people are able to assign a coreferential value to elements from different languages (such as *Mbeki se* and *I*). Language users can cross the border of their own language if the hearers may be assumed to share the frame of reference in which an expression from another language is well understood when used in a linguistic pattern of their own language.53 As a matter of fact, the same explanation can be considered to apply to monolingual compounds in which the modifier is a textual passage consisting of one or more “utterances”.

The third point concerns the nature of the relationship between the constituents of the compound. The question is not whether there is such a relationship, since a compound does not even exist without one. Rather, the question is “how is the head *toespraak* related to the modifier *I am an African*?” The semantics of *toespraak* allows for the idea that a speech is used to treat a theme. Therefore, the speech concerned could be referred to by means of a compound with *toespraak* “speech” as the head and the title of the speech, representing the speech’s content, as the modifier. However, in the case of *I am an African-toespraak*, we might assume that *I am an African* was not the title of the speech or its content in brief. What we can infer from (25) is that in his speech Mbeki focused on defining what it means to be an African. In other words, the speech at issue can be related recognizably to the modifier, i.e. the “utterance” *I am an*.

53. French *je* in the utterance below can be interpreted as exophorically related to the speaker (*pace* Bresnan & Mchombo 1995:194), although the compound *je ne sais quoi* feeling tends to be conventionalized (13 hits by Google at 31/05/2004; for *je ne sais quoi* quality, Google lists circa 100 hits at 31/05/2004).

So many of these songs have that *je ne sais quoi* feeling that falls somewhere between definite descriptions. (www.bluedark.com/reviews/Angelique_Kidjo–Black_Ivory_Soul.htm; my italics)
African, being a salient issue, also due to its being reminiscent of Kennedy’s *Ich bin ein Berliner* speech in 1961. However, in order to produce and to interpret the compound, the importance of Mbeki’s statement may be assumed to have been discussed before.

Utterance (24) is taken from an article starting with a composite motto of five identity statements, including *Ich bin ein Berliner* and *I am an African*. In the first paragraph of the article, the writer reminds the reader that the latter statement is a quote from the speech Mbeki held in 1996 “ter bevordering van ‘n Afrika-identiteit” “to enhance an Africa identity”. Utterance (24) turns up in the thirteenth paragraph of the sixteen-paragraph article. Since Mbeki’s speech is not mentioned in any of the eleven paragraphs in between, the writer needs a powerful reference to it in order to lay a foundation for the central issue in (25).

The fact that the writer can refer to Mbeki’s speech adequately by using the nonce compound *I am an African-toespraak* shows that knowledge established in the past can be activated and appealed to. It is, however, important to note that the compound *I am an African-toespraak* is chosen instead of, for example, the expression used in the first paragraph “toespraak gehou ter bevordering van ‘n Afrika-identiteit” “address given in order to enhance an Africa identity”. What can be deduced from the writer’s choice? To use the compound, both the writer and the reader have to reconceptualize the information given in the motto, the first paragraph, and the following paragraphs about language and identity. This information allows for the compact expression in the form of the nonce compound. The use of this compound is clearly agent-based, and understanding its actual referential value requires human conceptualizers.

Sometimes, the interpretation of a novel compound has to be postponed, for instance when the compound is used in the title of an article in which the explanation follows; see also (19). Consequently, the conceptualization of the current frame of reference has not finished or does not finish at the moment the morpheme to be interpreted turns up.

Langacker (1987:183–186) suggests describing the meaning of the word *arc*, for example, by relating it to the concept of a circle, or describing the meaning of the word *uncle* by relating it to the network of family relationships. Certainly, this type of background knowledge (see also Croft & Cruse 2004:Ch. 4) is important. However, what can be concluded from the use of a nonce expression such as *I am an African-toespraak* is that the actual conceptualization of the current frame of reference is the basis for using and understanding words, as is recognized by Langacker (1987:113–114).

---

54. The explanation of a novel compound is sometimes given afterwards; see, e.g., the following Dutch cases: *een ’sta-op-fauteuil’ voor een ’kontje’ uit de stoel* (*dV*, 07/08/03) ‘a get up armchair that gives one a leg up in order to rise from the chair’, and *een ik-ben-moe-beweging* (*wrijven in de ogen*) (*Volkskrant Magazine*, 04/10/03) ‘an I am tired gesture (rubbing over the eyes)’. The need for looking ahead is also discussed by Anstey (2006).
5. Concluding remarks

When communicating, speakers (writers) and hearers (readers) are assumed to conceptualize a frame of reference on the basis of the information that appears to be relevant at the current stage of the ongoing discourse. Speakers (writers) select the morphemes which are best suited to express their thoughts in an efficient and effective manner. I have attempted to show how possessive pronouns, the genitive, and novel compounds qualify for such a task. They can be used to refer to a wide range of cognizable relationships without the need for a detailed specification. The relationships which are compatible with possessive pronouns, the genitive, and novel compounds must correspond to what can be integrated into the conceptualization of the frame of reference being updated on the basis of one or more previous or subsequent utterances (or on the basis of the context of communication and knowledge of the world). The updating of the frame of reference related to the utterance cannot be processed by an automatic procedure, but has to be performed by conceptualizing human beings. Describing the semantico-pragmatic factors and procedures which serve to specify and determine speakers’ intentions, well-embedded in the current frame of reference, is in my view the central task of grammarians; it calls for a speaker/hearer-based grammar, i.e. a cognitive-pragmatic grammar focussing on utterances taken as linguistic acts, rather than as the results of linguistic acts.

References


Kilgariff, A. (2003). 'I don't believe in word senses.' In Nerlich et al. (Eds.), 361–391.


A speaker/hearer-based grammar


Index

A
abscene 175–179, 181–189
absolutive 38, 49, 143, 157
acquisition 68, 103, 137–142, 245
activation 68, 72, 76, 245, 247, 251, 285, 287, 294, 295, 297, 339, 346, 349
addressee’s meaning 354
adjective 13, 14, 23, 26, 28–33, 43–49, 52, 66, 67, 112, 156
adverb 3, 13, 23, 26, 30, 43
epistemic – 2, 9, 10, 15–18, 20, 22, 23, 24, 25, 30–33, 35–37
geographical 3, 46–51, 54–56
adverbial 3, 13, 35–38, 46, 56, 112, 119, 182, 209, 210–213, 222, 224, 300, 321; see also clause
adjetival 2, 9, 10, 15–18, 20, 22, 23, 24, 25, 30–33, 35–37
agent 62, 103, 105, 135, 142
143–155, 156, 157, 202, 218, 281, 285, 286, 322
agreement 107, 141, 143, 147, 150, 152, 153, 283, 316
algorithm 236, 275, 277, 285, 366
ambiguity 39, 43, 46, 47, 50, 51, 53, 54, 62, 73, 76, 96, 155
animacy 60, 61, 63, 70, 71, 76, 141, 146, 156
argument linking 99
aspect 108, 153, 163, 188, 189, 277, 279
nominal – 278, 280
auxiliary 20, 157, 185, 279, 283
modal – 2, 3, 10, 19, 22, 23, 32, 35, 37, 40, 43, 45, 53, 110, 159–161, 163–168, 170, 171, 251, 314, 328
bystander

B
benefactive 84, 90–92, 94, 96, 97, 99, 261–264
blend 43, 51, 52, 228
cognitive

B
Bsymbol

B
behaviourist 313, 321
case marking 143, 151, 152, 154, 155
causative 257, 259, 264, 269–272
clause

B
cognitive

B
decrease 138–140, 342–346, 348–350
ditransitive – 3, 83, 84, 88–91, 96
embedded – 200, 220
exclamative – 163
imperative – 29
main – 120, 127, 178, 210, 212, 218, 219, 222
relative – 115, 116, 143, 212, 219, 220, 233
simplex – 270–272
subordinate – 147, 212, 214, 375
that – 219
transitive – 272
C
case marking 143, 151, 152, 154
causative 257, 259, 264, 269–272

C
representational level 3, 210, 215–218, 223–225, 310, 333, 339, 344, 328
resultative 261–263, 265, 269
Role and Reference Grammar 1, 3, 142, 236, 248, 275, 304
S
schematicity 241, 252, 258, 272, 273
semantics 18, 40, 97, 175, 189, 229, 230, 234, 239, 241, 248, 262, 267, 275, 277, 281, 304, 356, 364, 365, 367, 369, 378
semantics-to-syntax linking algorithm 275, 285
sentence meaning 358, 366
spatial deixis 177
speech act 39, 43, 51, 54, 56, 162, 216, 217, 237, 238
speech production 310, 319, 337–343, 350
structural-functional grammar 304
structure-dependence 313, 315
subjectification 33, 165, 169
syntactic pattern 18, 23, 266, 276, 339
Systemic Functional Grammar 104, 197, 276
T
taxonomy 187, 234, 269
temporal 129, 143, 161, 183, 188, 194, 202, 203, 212, 215–220, 310, 318, 319, 343, 348
– deixis 123, 177
term, as used in Functional Grammar 193, 194, 195, 199, 202, 203–207, 213, 251, 314, 324
top-down 277, 310, 318, 322–324, 337, 338, 346, 349, 350, 358, 363
topic 68, 69, 72, 75, 79, 80, 103, 105–106, 110, 128, 129, 144, 145, 149, 175, 286, 288, 294, 296, 316, 321, 322, 324, 344
– continuity 2, 107, 123, 127–129, 134, 135
topicality 60, 63, 67–73, 76, 77, 80, 81, 103, 106, 128, 141, 145, 150, 156, 296
transitive 98, 148, 260, 262, 266, 269–272, 285
typology 97, 99, 100, 103, 141, 142, 177, 237
U
utterance meaning 354, 358, 359, 366
V
vagueness 54
variation 20, 23, 65, 85, 88, 89, 94, 95–100, 107, 110, 118, 124, 125, 128, 184
verbal noun 175, 176, 179, 180, 182, 184–189
vowel harmony 340
W
word order 64, 65, 67, 73, 118, 149, 160, 162, 187, 257, 296, 319, 320, 339
A complete list of titles in this series can be found on the publishers’ website, www.benjamins.com