Mapping Linguistic Data

Essays in honour of Liliane Haegeman

Metin Bağrıaçık, Anne Breitbarth, and Karen De Clercq (eds.)
Cover image: The White Cat (F. Marc)
Foreword

Liliane Haegeman’s oeuvre is characterised by the aim to advance syntactic theory based on strongly empirically informed research. A special interest of hers is the syntax and morphology of dialects, in particular (West) Flemish. At the same time, Liliane has worked together with an immense number of linguists on phenomena in an even larger number of languages. Her knack for puzzling data and rigid formal analysis have influenced the direction of research for many years. She has shaped (and continues to shape!) the field by her seminal work on verbal syntax, negation, adverbial clauses and main clause phenomena, and many generations of students have grown up on a healthy diet of several editions of the *Introduction to Government & Binding Theory* and, subsequently, *Thinking Syntactically*.

With this WebFestschrift, that contains linguistic and non-linguistic contributions, we want to thank Liliane for what she did for us and means to us as her students, colleagues and linguistic friends. We also want to wish her a very happy 65th birthday and a good retirement. Of course, we hope that Liliane will remain active in the field and that she will keep mapping linguistic data and astonish us all.

All the best to you, Liliane, and a very happy birthday!

Metin Bağrıaçık
Anne Breitbarth
Karen De Clercq
# Table of contents

Possessors, foci and topics in the Greek DP  
Artemis Alexiadou & Melita Stavrou .................................................. 8

A negative concord puzzle tied to French exceptives  
J.-Marc Authier ................................................................. 15

Subject relative clauses without a complementizer in a Modern Greek dialect  
Metin Bağriaçık ............................................................... 25

Subjunctive selection in French vs. Balkan  
Lena Baunaz & Eric Lander .......................................................... 35

‘U is’ or ‘U bent’? Polite forms of address and verbal agreement in historical Dutch and Ibero-Romance  
Miriam Bouzouita & Anne Breitbarth .................................................. 50

The Haegeman test: A’-movement in the middle field  
Hans Broekhuis ................................................................. 60

The Root where it should not be. On internal argument drop sentences in Italian  
Carlo Cecchetto ................................................................. 68

A Note on Relative Clauses with Split Antecedents  
Guglielmo Cinque ................................................................. 75

A dummy and a diminutive in Dutch verbal lore  
Norbert Corver ................................................................. 82
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>De kunst van de behoedzaamheid / The art of precaution</td>
<td>97</td>
</tr>
<tr>
<td>Karen De Clercq</td>
<td></td>
</tr>
<tr>
<td>'What would you rather me say?'</td>
<td>99</td>
</tr>
<tr>
<td>Marcel Den Dikken</td>
<td></td>
</tr>
<tr>
<td>While we’re on the Subject...</td>
<td>105</td>
</tr>
<tr>
<td>Nigel Duffield</td>
<td></td>
</tr>
<tr>
<td>Some thoughts on the early Dutch null subject</td>
<td>114</td>
</tr>
<tr>
<td>Maria Teresa Guasti</td>
<td></td>
</tr>
<tr>
<td>Negative concord and Old English clause structure:</td>
<td>119</td>
</tr>
<tr>
<td>some afterthoughts</td>
<td></td>
</tr>
<tr>
<td>Eric Haeberli</td>
<td></td>
</tr>
<tr>
<td>Mapping the speaker’s mind</td>
<td>129</td>
</tr>
<tr>
<td>Virginia Hill</td>
<td></td>
</tr>
<tr>
<td>Some information-structure properties of Negative Preposing</td>
<td>136</td>
</tr>
<tr>
<td>in English and Spanish</td>
<td></td>
</tr>
<tr>
<td>Ángel L. Jiménez-Fernández</td>
<td></td>
</tr>
<tr>
<td>Negative adjectives</td>
<td>146</td>
</tr>
<tr>
<td>Wim Klooster</td>
<td></td>
</tr>
<tr>
<td>Fanbrief / Fan mail</td>
<td>154</td>
</tr>
<tr>
<td>Karen Lahousse</td>
<td></td>
</tr>
<tr>
<td>Morphology Driven by Syntax: XP Movement and Allomorphy</td>
<td>161</td>
</tr>
<tr>
<td>Tom Leu</td>
<td></td>
</tr>
</tbody>
</table>
Multilingualism as the new comparative syntax
Terje Lohndal ................................................................. 171

To Liliane, with appreciation
Shigeru Miyagawa ........................................................... 179

Hoe cool is dat wel niet! How pseudo-questions and expletive negation in Dutch
Rachel Nye and Lieven Danckaert ..................................... 182

Polarity and other distributional properties of uitkijken / opletten / oppassen ('look out')
Albert Oosterhof ............................................................ 200

The absence of classifiers in numeral classifier constructions in Vietnamese
Trang Phan ................................................................. 208

Embedded interrogatives as free relatives
Cecilia Poletto & Emanuela Sanfelici ................................ 214

For Liliane: Friend, muse, linguist extraordinaire, and world’s second most famous Belgian
Andrew Radford ............................................................. 226

Left peripheral NEG as a discourse particle
Henk van Riemsdijk ......................................................... 234

‘Recycling’ evidentiality: a research program
Johan Rooryck ............................................................. 242
An emerging veridical complementizer to čto in Russian
Alexandra Simonenko & Anastasiia Voznesenskaia ........................................ 262

Clause-initial subject doubling in Dutch dialects
(Or: Liliane was right after all)
Jeroen Van Craenenbroeck & Marjo van Koppen ........................................ 269

Here, there, and (every)where
Guido Vanden Wyngaerd .......................................................... 280

Grammaticalisation processes in Flemish Sign Language
Mieke Van Herreweghe .......................................................... 287

“Stemmen uit het verleden”: Het transcriptieproject bij de Vakgroep Taalkunde—Nederlands van de Gentse Universiteit
Jacques Van Keymeulen & Anne-Sophie Ghyselen ........................................ 296

Diary null subjects: an analogy with imperatives?
Andrew Weir .......................................................... 306

Recording and explaining: exploring the German ditransitive alternation
Klaas Willems, Ludovic De Cuypere & Hilde De Vaere ............................... 313
Possessors, foci and topics in the Greek DP

Artemis Alexiadou & Melita Stavrou

1 Introduction

The position(s) and the status of the DP initial possessive genitive in Greek have been extensively discussed in the literature since the mid/late eighties. Although not the principal focus of investigation in Horrocks & Stavrou’s work (1987), the nature of possessive genitive constituted their main argument for the existence of a functional category above NP, namely DP, in Greek.

In this squib, dedicated with love and respect to Liliane and her extensive work on possessors in West Flemish (and cross-linguistically), we return to the issue of the position and function of the possessive genitive and look with more detail at the preposed genitive and its possible interpretations.

The squib is structured as follows. First, we briefly summarize the basic facts and argumentation of Horrocks & Stavrou (1987). The main points of this outline concern the parallelisms between D and I/C in the sentential domain and the focus interpretation that both possessors and interrogative words get when moved to the specifier position of their respective functional categories. We also relate the focus interpretation with the (im)possibility of the genitive being doubled by a pronominal clitic inside the DP. Building on an idea first encountered in Giusti & Stavrou (2008), we show that the pronominal (or preposed) possessor in Greek can only be stressed as focus and cases of apparent violation of that generalization which have been brought up over the years (see Giusti & Stavrou 2008) are instances of either topicalization or of dislocation (as Hanging Topic) of the whole DP that contains the possessor. Closing our squib, we summarize the functions and interpretations of the preposed genitive in Greek.
2 Horrocks and Stavrou (1987)

The point of departure of Horrocks & Stavrou were nominal phrases where the possessor, in the form of a full DP whose lexical head (as well as other possible material) bears morphological genitive case, is found in a position preceding the article:

(1)  
   a. tu fititi to vivlio  
       the.GEN.SG student.GEN.SG the book  
       ‘the student’s book’
   b. ton pedjion to domatio  
       the.GEN.PL kids.GEN.PL the room  
       ‘the kids’ room’

The position of the genitive in (1) is marked, since the intonation and interpretation it gets is that of a (contrastive) focus, in contrast to the neutral interpretation it gets in its ‘base’ position, following the noun. The focus function of the preposed genitive automatically means that the possessor cannot be doubled by a clitic:

(2)  
   a. *Tu fititi to vivlio tu katastrafike.  
       the.GEN.SG student.GEN.SG the book CL.3SG.GEN was destroyed  
       ‘The student’s book was destroyed.’
   b. *ton pedjion to domatio tus  
       the.GEN.PL kids.GEN.PL the room CL.3PL.GEN

Horrocks & Stavrou (1987) put special emphasis on the fact that the same distributional pattern as in (1) is observed in interrogative DPs, in which the interrogative genitive also precedes the (definite) article (for more data and detailed discussion see Horrocks & Stavrou 1987):

(3)  
   tinos to aftokinito?  
   whose the car (from: to aftokinito tinos? ‘the car whose’)  
   ‘Whose car?’

There is a further eloquent parallelism between interrogative clauses and inter-

---

1 It must be noted that clitic doubling (of the possessor) is not attested inside the DP irrespectively of the position of the possessive genitive (Alexiadou & Stavrou 2000, Giusti & Stavrou 2008). The fact that it is predictably disallowed with focused genitives is independent and explained on the basis of the focus character of the genitive.
rogative DPs:

(4) Ti ekane?
    what did.3SG
    ‘What did he do?’

is exactly parallel to (3), the former showing fronting of the interrogative pronoun in the sentence, the second in the nominal expression. This parallelism led Horrocks & Stavrou to assume that within the nominal phrase there must be a grammatical position for the interrogative possessive pronoun to land to (internal wh-movement) parallel to CP in clauses, where wh-constituents move to. Such a position, a specifier position by analogy to Spec,CP, must be in front of the article, given the word order in (3). In that case, the article is the head of the projection at the specifier of which constituents from a position lower in the nominal phrase move to. Horrocks & Stavrou (1987) named that projection Art(icle)P. It was the same projection that was labelled DP in Abney (1987). The crucial fact underlined by Horrocks & Stavrou was that the pre-article position is an A′-position, similar to the Spec, CP. Thus, in Greek, in contrast to English, the same relationship holds between DP and NP as between CP and VP. In both cases, the specifier position of the functional layer is an A′-position and may host constituents moved from within the lexical category they dominate (operator movement); in English, DP is parallel to IP. The status of Spec,DP as an A′-position is used in Alexiadou (2016) to explain the unavailability of possessor doubling with full DPs in Greek. Giusti & Stavrou (2008) note that in certain cases involving pronominal possessors, the possessor can be clitic-doubled (see (5)). This is impossible when the possessor is a full DP (6):

(5) To vivlio mu emena den pulithike katholu.
    the book CL.1SG me.GEN STR not sold at all
    ‘My book was not sold at all.’

(6) Petaxa to vivlio-(*tu) tu fititi.
    threw.1SG the book-(*his) the student.GEN
    ‘I threw away the student’s book.’

German and other Germanic languages have possessor doubling with full DPs, as shown in (7):

(7) dem Vater seine Katz
    the father.DAT his cat
According to Alexiadou (2016), Greek differs from German(ic), which has possessor doubling, in that in Greek, the pre-determiner possessor occupies an A'-position, as argued by Horrocks & Stavrou (1987). By contrast, in German(ic), the nominal possessor can be argued to occupy an argument position, as it receives no particular focal or contrastive stress, see Corver (1990), Haegeman (2004), and Georgi & Salzmann (2011) among others. The latter type of possessor can be doubled by a clitic, which Alexiadou assumes, following Corver, that occupies D0. This is not possible in Greek. In Greek, the possessive clitic is enclitic to the head noun and never realizes D0. Evidence for this comes from the observation that in German(ic) possessor doubling constructions, the pronoun inflects like a determiner and not like an adjective, the possessive pronoun determines the inflection of a following adjective, and the possessive pronoun is in complementary distribution with other determiners.

3 Possessors, foci, topics and the DP

However, Giusti & Stavrou (2008) make the following observation, which appears to weaken the generalization that the preposed possessor in Greek is always (contrastively) focused. There are instances of DP initial genitives that are not focalized but look like (contrastive) topics (although such instances are much rarer than the cases where the preposed genitive is focused):

(8) Tis Marias o petheros epathe egefaliko.
    the.GEN SG Maria.GEN.SG the father-in-law underwent stroke
    ‘Mary’s father-in-law underwent a stroke.’

In (8) the preposed genitive is not focalized and does not get emphatic/contrastive stress. The reading of (8) is something like ‘talking of Mary/As for Mary her...’.

Let us further consider the following context.

Context: Several of my former classmates face with health and other problems in their families; in particular:

(8) Tis Marias o petheros...

(9) Tis Elenis o jos epathe atihima sti thalassa.
    the.GEN SG Helen.GEN.SG the son had accident in the sea
    ‘Helen’s son had an accident in the sea.’
(10) Tu Kosta o aderfos horise prin dio mines.
    the.GEN.SG Kosta.GEN.SG the brother divorced before two months
    ‘Kosta’s brother got divorced two months ago.’

In these cases, the preposed possessor is not the focus but has rather the flavor
of a topic, probably a contrastive topic. Nonetheless, the clitic is not allowed in
such cases either:

(11) *Tis Elenis o jos tis epathe atihima sti
    the.GEN.SG Helen.GEN.SG the son CL.GEN.SG had accident in the
    thalassa.
    sea

Giusti & Stavrou (2008) hypothesize that rather than saying that the genitive
stands for a topic inside the DP, it is the entire DP that has this function, not just
the possessor. The topic character of the whole DP in cases such as [8]–[10] is
supported by the fact that it can be clitic doubled just like any other topicalized
object in Greek:

(12) Tis Marias ton pethero ton skotosan i
    the.GEN.SG Maria.GEN.SG the father-in-law CL.ACC.SG killed the
    jermani ston polemo.
    Germans in the war
    ‘Maria’s father-in-law was killed by the Germans during the war.’

In isolation, the possessor may only be focused – the across-the-board case as
said above. If this line of thought is along the right track, it means that the
preposed genitive in [8] was originally focused within the DP, but when the DP
moved to a topic position in the clause, the focus status of tis Marias was can-
celled or overridden by the topic status of the whole DP. Full assessment of non-
focused (preposed) genitives as in [8]–[10] requires a detailed examination of the
interaction of foci and topics in the clause and in the DP.

Let us next consider cases similar to those in [8]–[10] but with a clear pause
after the preposed genitive:

(13) Tis Marias, i aderñi *(tis) exi megalo provlima
    the.GEN.SG Maria.GEN.SG the sister CL.GEN.SG has big problem
    me ton antra tis.
    with the husband hers

12
‘Mary’s sister has a big problem with her husband.’ (Lit. ‘Let’s talk of Mary – she has...’).

In (13) the clitic inside the clause is required, in contrast with the previous case. Giusti & Stavrou assume that the possessor occupies a topic position in the sentence, not within the DP. If we take into consideration Anagnostopoulou’s (1997) distinction between Clitic Left Dislocation and Hanging Topic, the question is whether in (13) the extraposed genitive DP is an instance of a hanging topic or (left) dislocation. As commented by Giusti & Stavrou (2008), both dislocation and HT represent exclusively oral (and often casual) speech, hence it is difficult to obtain clear judgments from the speakers. Nonetheless, here we follow Giusti & Stavrou (2008: 419–421) and assume that (13) instantiates a case of dislocation. There are two reasons for this: first, the preposed genitive may not carry default nominative, something that is usual with nominal hanging topics; second, the genitive may not be co-indexed with an epithet or a demonstrative lower in the clause, something, again, typical of hanging topics. In (13) the genitive is adjoined (to the left, but it can also be adjoined to the right) to the IP (or CP) and from there it forms a chain with the clitic – an operator – that appears within the clause. The two share the same phi-features and (genitive) case.

In the light of the above discussion, we conclude that the possessive DP that is found before the definite article in the Greek DP is a focused constituent which has moved to Spec, DP attracted by the [+foc] feature on D. This movement parallels the movement of wh-constituents in the clause. Apparent violations of this generalization are instances of either topicalization of the whole DP that contains the preposed possessor to a sentential topic position, or of dislocation of the possessor as Hanging Topic (along the lines of Anagnostopoulou 1997) to IP or CP.

4 Conclusions

In Greek, Spec, DP is A′-position. It can host both foci, but not topics, coming from a lower position in the DP. In that case it bears a [+foc] feature. In the absence of such a feature, the possessor stays in its original merge position. When the preposed genitive DP is not focused, then two things can happen: either it is the entire DP that has moved to a sentential topic position, most likely as a contrastive topic, (if it is an object of a verb it can be clitic doubled) after the possessor was moved to the spec, DP inside the DP. Alternatively, the possessor is
dislocated and adjoined to the IP (or the CP) with a clear pause separating it from the rest of the clause (doubling is then mandatory).

References

Georgi, D. & Salzmann, M. 2011. DP-internal double agreement is not double agree: Consequences of agree-based case assignment within DP. *Lingua* 121. 2069–2088.
A negative concord puzzle
tied to French exceptives

J.-Marc Authier

1 Introduction

Elaborating on ideas put forth in Baciu (1978), a number of researchers such as O’Neill (2011), Homer (2015) and Authier (2018) have converged on the idea that French (ne) ... que exceptives like (1) are hidden quantificational comparatives that contain phonologically unrealized material. This material (capitalized in (1)) consists of a silent n-word rien ‘nothing’ and, as argued in Authier (2018), a silent de plus ‘more’.

\[(1) \text{ Elle (n’) aime RIEN DE PLUS que les perles.} \]

‘She only likes pearls.’

The evidence in favor of assuming the presence of the silent n-word RIEN ‘nothing’ is quite robust. Besides the fact that it can appear overtly in many varieties of French, including colloquial standard French, Homer (2015) observes that when an exceptive and an n-word are clause-mates, as in (2), they give rise to either a negative concord (NC) reading or a double negation (DN) reading.\(^1\) This suggests that (ne)...que exceptives do indeed contain a covert n-word.\(^2\)

\[^1\text{As has been noted (cf. Corblin (1996: 251), stress on the first (or the second) n-word seems necessary to bring out the double-negative reading.} \]

\[^2\text{Additional evidence comes from the fact that an exceptive adverb like seulement ‘only’, not being an n-word, does not trigger negative concord readings. Thus, (i) contrasts with (2) in that only (i) allows NC.} \]

(i) \text{ Personne ne porte seulement du bleu.}  

DN reading only: Everybody wears something besides blue.
As for the silent DE PLUS component, it too can appear overtly as long as overt rien is used as well. This is illustrated in (3a)-(3b).

(3) a. On (n’) a vu rien de plus que quelques phoques.
   ‘We only saw a few seals.’

b. *On (n’) a vu RIEN de plus que quelques phoques.

Thus, a reasonable hypothesis seems to be that (4a) is the maximal phonological realization of a quantificational comparative which can also be spelled out as the partial realization in (4b) and the minimal realization in (4c). That is, (4a)-(4c) are syntactically and semantically equivalent and differ only in terms of those features that are accessed by the phonological component.

(4) a. Je (n’) ai acheté rien de plus que des tomates.
   ‘I only bought tomatoes.’

b. Je (n’)ai acheté rien DE PLUS que des tomates.

c. Je (n’)ai acheté RIEN DE PLUS que des tomates.

This hypothesis immediately raises the question of whether the silent elements represented in capital letters in (4b) and (4c) come to be silent in the same way as lexical elements whose phonetic matrix is deleted at Spell-Out, instantiating the phenomenon known as ellipsis (see e.g. Merchant 2001, 2004). To answer this question, let us take as a point of departure the feature-based taxonomy of lexical items given in Her & Tsai (2015). This taxonomy assumes that canonical lexical items have formal features (FF), which are accessible in the course of the narrow-syntactic derivation, as well as phonological features (PFF) and semantic features (LFF). While all lexical items active in syntax must have FF, non-canonical lexical items may lack PFF, LFF, or both. For example, overt expletives have no LFF, base-generated silent elements like PRO and pro have no PFF and null expletives have neither PFF nor LFF. Further, lexical items with no PFF, which we will simply call silent elements (SEs), differ from elements whose silence is due to ellipsis in that while the former have no PFF to begin with, the latter enter the derivation with PFF (i.e. enter the derivation as canonical lex-
ical items) but their PFF are made invisible by ellipsis at the syntax-phonology interface (so-called PF deletion). A second important difference between SEs and elided elements has to do with the way in which they are subject to recoverability. That is, the meaning of SEs is recoverable from their pronounced counterparts, which means that SE do not require overt antecedents. PF-deleted elements, on the other hand, are recoverable through overt discourse antecedents. Given this taxonomy, the silent components in (4a)-(4c) must be seen as SEs rather than elided elements given that they are interpreted not via a discourse antecedent but, rather, by accessing the meaning of their PFF-endowed counterparts in the lexicon. Keeping these characteristics in mind, I will devote the remainder of this squib to the issue of availability of partial and minimal exceptive ne...que realizations under prepositions. As I will show, while their syntactic derivation is consistent with the hypotheses formulated above, their semantic behavior as regards the availability of negative concord constitutes an unexpected puzzle which I will carefully lay out but for which I will offer but speculative remarks.

2 Deriving prepositional exceptives

Based on paradigms like (5), it has been widely assumed in the literature (see e.g. Gross 1977: 90) that exceptive que can never follow a preposition.

(5) Je (ne) compte que sur/*sur que son intégrité.
   I NEG count than on/*on than his integrity
   ‘I count only on his integrity.’

However, some have noted (e.g. Damourette & Pichon 1943: 220 and Gaatone 1999: 105) that this characterization does not always hold in colloquial registers, as (6) illustrates.

(6) a. Faites trois séances si vous voulez, moi je ne viendrai à que
do three sessions if you wish me I NEG will-come to than
two
   ‘Organize three sessions if you wish; me, I’ll come to only two of
   them.’
   (M. ABA, July 2, 1919, recorded in Damourette & Pichon 1943: 220)
b. [...] des ministres **avec que** des vraies factures [...] some ministers with **than** some genuine bills

‘[...] ministers with only genuine bills’

(TF1, May 13, 1991, recorded in Gaatone 1999: 105)

On closer inspection, the full prepositional exceptive paradigm turns out to be comprised of three distinct cases. First, there are instances, illustrated in (7), in which the que in minimal realizations not only may but must follow the preposition regardless of register. These typically involve PPs complement to N and do not allow (optional) *ne*.

(7) a. Ils ont un menu **avec que** des produits bio.

they have a menu with **than** some products organic

‘They have a menu with only organic products.’

b. *Ils ont un menu **qu’avec** des produits bio.

Under the view adopted here, the maximal realization for a sentence like (7a) is as in (8).

(8) Ils ont un menu **avec rien de plus que** des produits bio.

The minimal realization and the partial realization of (8) are then obtained by using the silent counterparts to the overt elements in bold, yielding (9a)–(9b).

Further, given the source in (8), there is no way to derive the ungrammatical (9c), in which que precedes the preposition.

(9) a. Ils ont un menu avec RIEN DE PLUS que des produits bio.

b. Ils ont un menu avec rien DE PLUS que des produits bio.

c. *Ils ont un menu qu’avec des produits bio.

Next, there are maximal realizations, such as the one in (10), in which the preposition introducing the phrase that is the object of the comparison must be repeated.

(10) Ils (ne) causent de rien **de plus que** *(de)* politique. they NEG talk of nothing of more than *(of) politics ‘They only talk about politics.’

Both the partial and minimal realizations of such sentences involve the overt realization of at least one of the two instances of the preposition. The first instance of the preposition can only be overt if rien is overt. Further, if only one instance of the preposition is overt, it must be the second one. This correctly rules in (11a)-
and rules out (11d)–(11f).

(11)  
  a. Ils (ne) causent de rien DE PLUS que de politique.
  b. *Ils (ne) causent de RIEN DE PLUS que de politique.
  c. *Ils (ne) causent de rien DE PLUS que de politique.
  d. *Ils (ne) causent de RIEN DE PLUS que de politique.
  e. *Ils (ne) causent de rien DE PLUS que de politique.
  f. *Ils (ne) causent de RIEN DE PLUS que de politique.

We are, however, left with the following questions, for which I have no answers at this time: (a) why is the overtness/presence of the first instance of the preposition contingent upon rien being overt; (b) why must it be the second instance of the preposition that is overt when only one instance of the preposition is overt/present; (c) is the first instance of the preposition, crossed out in the representations in (11a)–(11b), simply omitted or is it syntactically present but phonologically silent and why?

The third and final set of facts consists of maximal realizations like (12a)–(12b), which feature one or two instances of the preposition being used.

(12)  
  a. Je (n’) ai cuisiné avec rien de plus que des produits frais.
       I NEG have cooked with nothing more than some fresh ingredients.
       ‘I cooked with only fresh ingredients.’
  b. Je (n’) ai cuisiné avec rien de plus que des produits frais.

We thus have two possible sources for partial and minimal realizations. Stemming from the source in (12a), which contains only one instance of the preposition, are the partial and minimal realizations in (13), both of which are attested.

(13)  
  a. Je (n’) ai cuisiné avec rien DE PLUS que des produits frais.
  b. Je (n’) ai cuisiné avec RIEN DE PLUS que des produits frais.

The source in (12b), which contains two instances of the preposition, additionally allows us to generate the partial and minimal realizations in (14a)–(14d), which are attested as well. Note, however, that, based on the data in (14), one must again assume, as we did for (11), that the first instance of the preposition can only be overt if rien is overt, i.e., not a SE (or a trace – cf. (14d) which differs from the partial realization in (14b) only in that rien appears in pre-participial position, a case of so-called quantification at a distance). Further, if only one instance of
the preposition is overt, it must be the second one.

(14) a. ??Je (n’)ai cuisiné avec rien DE PLUS qu’avec des produits frais.
b. Je (n’)ai cuisiné avec rien DE PLUS qu’avec des produits frais.
c. Je (n’)ai cuisiné avec RIEN DE PLUS qu’avec des produits frais.
d. Je (n’)ai rien, cuisiné avec t, DE PLUS qu’avec des produits frais.
e. *Je (n’)ai cuisiné avec RIEN DE PLUS qu’avec des produits frais.

In sum, despite some needed stipulations concerning which instances of the preposition may or may not be overt, our general assumptions concerning the nature of the full realization of both partial and minimal realizations allow us to generate all of the attested syntactic realizations of these constructions. In what follows, however, I will show that things are not as straightforward at the syntax-semantics interface. Specifically, I will demonstrate that some partial and minimal realizations unexpectedly exhibit a behavior that diverge from that of their full realizations when it comes to the availability of negative concord readings.

3 A negative concord puzzle

As is well-known, in French, when two n-words are clause-mates (cf. Déprez 1999 among many others), they may (but need not) give rise to a NC reading. Thus, a sentence like (15a), which contains the two n-words jamais ‘never’ and personne ‘nobody’, is ambiguous between a NC (single negation) reading (15b) and a DN (double negation) reading (15c).

(15) a. Il n’y a jamais personne sur cette plage.
    there is never nobody on this beach
    (NC interpretation)

(16) a. Je (n’) ai jamais cuisiné avec rien de plus qu’avec des produits frais.
    I NEG have never cooked with nothing of more than-with some ingredients fresh

The fact that the two maximal realizations in (16a)-(16b), which contain the two n-words jamais ‘never’ and rien ‘nothing’, are reported by a majority of speakers to be similarly ambiguous between (16c) and (16d) immediately suggests that the two n-words are clause-mates.
b. Je (n’)ai jamais cuisiné avec rien de plus que des produits frais.
c. I’ve never cooked with anything more than fresh ingredients. (NC reading)
d. I’ve always cooked with something more than fresh ingredients. (DN reading)

This being the case, we then expect that adding the n-word jamais to the five partial and minimal realizations in (13) and (14), as in (17) and (18), should yield a similar ambiguity between a NC and a DN reading. This follows from our assumptions that (a) partial and minimal realizations are syntactically indistinguishable from their maximal realization sources and (b) the SEs they contain are semantically recoverable by accessing the meaning of their PFF-endowed counterparts found in their maximal realizations. This prediction, however, turns out to be incorrect. Although there is some variation in native speakers’ judgments, the great majority of my informants agreed on the following interpretations.

(17)  
  a. Je (n’)ai jamais cuisiné avec rien DE PLUS que des produits frais.  
      Unambiguous: DN reading only
  b. Je (n’)ai jamais cuisiné avec RIEN DE PLUS que des produits frais.  
      Unambiguous: DN reading only

(18)  
  a. Je (n’)ai jamais cuisiné avec rien DE PLUS qu’avec des produits frais.  
      Ambiguous: NC and DN readings possible (DN reading preferred)
  b. Je (n’)ai jamais cuisiné avec RIEN DE PLUS qu’avec des produits frais.  
      Ambiguous: NC and DN readings possible (NC reading preferred)
  c. Je (n’)ai jamais RIEN i cuisiné avec t i DE PLUS qu’avec des produits frais.  
      Ambiguous: NC and DN readings possible (NC reading preferred)

While the interpretations tied to the partial and minimal realizations in (18) linked to the maximal realization source in (16a) do conform to our predictions, those in (17), which stem from the maximal realization in (16b) do not. Specifically, while (16b) allows NC, its purported partial (17a) and minimal (12b) realizations do not. Descriptively, the presence of one or more SEs preceded by an overt instance of the preposition somehow blocks NC. Explaining why this is so is not immediately obvious. It seems clear, however, that what contributes to the unavailability of NC in (17) is the combination of two factors; namely the use of SEs and the fact that the quantificational comparative is c-commanded by a phonologically overt preposition.
Regarding the first factor, one could conjecture that in particular syntactic configurations, the use of phonologically unrealized material blocks NC readings. This is not an implausible assumption, given that a case can be made for ellipsis having such a blocking effect on NC. To explain, Merchant (2004) has argued, based on a range of connectivity effects, that fragment answers are derived from full sentential structures, subject to ellipsis. If this is correct, then (19) should be assumed to have the representation in (19), where the crossed-out material represents syntactically present material that has undergone PF-deletion.

(19)  A: Qui n’a jamais pleuré? ‘Who’s never cried?’
     B: Personne. ‘Nobody.’
     C: Personne n’a jamais pleuré.
     D: Personne n’a jamais pleuré.

As pointed out in Corblin (1996: 251), while a non-elliptical answer like (19) is ambiguous between a NC and a DN reading, its corresponding fragment/elliptical answer in (19) cannot have a NC reading but must be interpreted in a bi-negative fashion (i.e., as meaning “Everybody has cried.”). Thus, fragment answer ellipsis is one environment in which the presence of phonologically silent material negatively affects the availability of NC.

In the case under consideration in (17), however, saying that the presence of SEs blocks NC is insufficient since failure to license NC only occurs if the SE in question is c-commanded by an overt preposition (or, more simply, a preposition, if we assume that the crossed-out prepositions in cuisine6 are not syntactically represented). This has the flavor of a locality condition but, unfortunately, locality conditions on NC are poorly understood and have yet to be spelled out beyond the widespread observation that NC appears to be clause bound. Indeed, while it seems accurate to assume that French n-words that are not clause-mates fail to yield NC readings, it can be shown that n-words that are clause-mates do not always participate in NC. For example, n-words complement to the preposition pour ‘for’ participate in NC with other clause-mate n-words when pour is taken to mean ‘in favor of’ (20a) or introduces the stimulus of a psych predicate (20b), but do not participate in NC when pour expresses a reason (20c).

(20)  a. Personne ne s’est prononcé pour rien. (ambiguous)
       Nobody claimed to be in favor of anything. (NC)
       Everybody claimed to be in favor of something. (DN)
    b. Personne ne s’est inquiété pour rien. (ambiguous)
Nobody worried about anything. (NC)
Everybody worried about something. (DN)

c. Personne n'a été puni pour rien. (unambiguous)
Nobody got punished for anything. (NC)
Everybody got punished for something. (DN)

These facts, which, to the best of my knowledge, have never been discussed in the literature, would force proponents of the clause-bound characterization of the locality conditions on NC to attribute a clausal structure to the complement of pour in (20c), but not in (20a)-(20b), a dubious move at best. A possibly more straightforward interpretation of (20) is that the set of interpretative options normally available to an n-word complement to a preposition can be reduced by the sort of thematic role that preposition assigns to it.

The thematic properties of prepositions have also been shown by Authier (2016) to play a role in the licensing of the SE pro in the context of orphan prepositions. For example, as shown in (21), the preposition dedans 'in' can license a pro with definite interpretation if it takes a complement that denotes a material entity with well-defined boundaries (e.g., a package – cf. (21a)) but not if it takes a complement that denotes a spatial entity with ill-defined or unknown boundaries (e.g., the streets of Montmartre – cf. (21b)).

(21) a. Ce colis, il y a un cadeau dedans pro.
   this package there is a present in (it)

   b. *Les rues de Montmartre, elle vend des crêpes dedans pro.
   the streets of Montmartre she sells crepes in (them)

Returning to the interpretive contrast between the full realization of the prepositional exceptive in (16b) and its partial and minimal realizations in (17), we can now reinterpret the facts in a similar light. That is, while the preposition avec 'with' licenses both the overt and SE versions of the exceptive, the latter are more restricted in their interpretation or, to put it slightly differently, although SEs are recoverable from their pronounced counterparts, they display a more restricted set of interpretations when they are complement to a preposition. This is, of course, rather speculative, but it does suggest that there might be a link between the semantic properties of SEs and the selectional restrictions of the prepositions that license them. I will leave this as an open question for future research.
References

This short note is written in gratitude for Liliane Haegeman’s immense contributions to syntactic theory and to the life of many young scholars, including me, in so many ways. I hope that these few pages are worthy of her and of her insights into the nature of many phenomena—among them subject contact relatives.

1 Introduction

Restrictive subject and object relative clauses (hereafter RCs) in Pharasiot Greek (hereafter PhG) are finite clauses that are, in the unmarked case, introduced by the complementizer tu ‘that’. The head, external to the RC, can linearly precede and follow the RC (Bagriacik & Danckaert 2018).

(1) a. Ídha [an gorítsi [tu kathéti ačí monaxó ts]].
   saw.sg a girl that sit.3sg there on-her-own
b. Ídha [tu kathéti ačí monaxó ts] an gorítsi.
   saw.sg that sit.3sg there on-her-own a girl
   ‘I saw a girl (who/that) was sitting there on her own.’

Andriotis (1948: 51) and Anastasiadis (1976: 248) claim that there exists a third, marked, option for forming restrictive relative clauses in PhG: “[RCs] in Pharasiot Greek, especially the restrictive ones, are occasionally introduced into
the discourse without a [...] relative adverb [my translation]” (Anastasiadis 1976: 248). These authors illustrate both subject and object RCs without tu ‘that’. In this note, I focus only on subject RCs. Object RCs without tu ‘that’ merit a proper study on their own. (2) is an example of tu-less subject RC provided by Andriotis (1948: 51).

(2) Ídha [an goritsi kathéti aćí monaxó ts].
    saw.sg a girl sit.3sg there on-her-own
    ‘I saw a girl (who/that) was sitting there on her own.’

The aim of this note is to assess whether or not structures as (2), which lack tu ‘that’, do indeed qualify as (some sort of) RCs. To this effect, I compare them to English subject RCs with the null variant of the relative pronoun or with no overt complementizer. I conclude that the existence of complementizerless subject RCs in PhG should be recognized, verifying the claim by Andriotis (1948) and Anastasiadis (1976).

In the next section, I briefly introduce English subject RCs that do not feature an overt relative pronoun or a complementizer, and further summarize two main analyses from a generative perspective.

2 Subject RCs without overt relativizers in English

In informal or colloquial speech, English allows subject RCs in which the relative pronoun or the complementizer can be omitted in certain syntactic environments. Such RCs are dubbed by Jespersen (1928: 143ff) ‘subject contact clauses’ and they are widely known as ‘subject contact relatives’ today (Doherty 1994, 2000, Henry 1995; hereafter SCRs). As opposed to regular subject RCs, the distribution of SCRs is reported to be limited: some of the typical contexts in which they are allowed are (i) existential copular sentences (3a), (ii) existential have sentences (3b), and complements of the predicates know, meet and invent, when these complements introduce a new referent into the discourse (3c) (Doherty 2000: 72, Henry 1995: 125).

According to Anastasiadis (1976) tu is an indeclinable relative adverb, comparable to the Modern Greek complementizer pu ‘that’. In the rest of this note, I will refer to tu as a complementizer. See Bağrıaçık (2018) for discussion of the environments in which tu is used.

Doherty (1994) argued that in order for a SCR to be licensed, “[the] noun phrase modified by [it] must be interpreted as non-referential.” See, however, Doherty (2000), where he no longer maintains this and leaves the question of the precise conditions for SCR-licensing open.
a. There's something keeps upsetting him.
b. I have this friend lives in Dublin.
c. I know a smart Greek owns maybe twenty restaurants.

SCRs are claimed to be excluded from other grammatical positions, such as subject position (4a) or the indirect object position (4b) (Doherty 2000: 72-73, Henry 1995: 125-126).

(4) a. *The man worked there was a friend of mine.
b. *I gave a ticket to a man comes every day.

SCRs have hitherto received two main types of analysis in the generative literature. According to one, which is advanced by Doherty (1994, 2000), SCRs are a variant of true restrictive RCs. They differ from regular RCs in one major respect: whereas regular RCs are standardly assumed to involve a full CP, SCRs have a reduced internal structure; more specifically, they lack the CP layer. According to this account, the example (3a) has the (simplified) structure in (5).

(5) There's [DP [NP something [TP e keeps upsetting him]]].

The reader is referred to Doherty (1994, 2000: 81-87) for the details of his analysis, which are immaterial for the current note. The important point is that this approach claims that an SCR is a genuine subordinate clause.

According to a second line of analysis, SCRs involve a topic-comment structure (Henry 1995: 131-135, den Dikken 2005). As a corrolary of this, what superficially is a matrix clause hosting the SCR is analyzed as a topic clause, whose function is to identify one of its subconstituents, the head nominal in particular, as a focus. In turn, what is taken to be a subordinate clause in Doherty’s (1994, 2000) analysis is identified by Henry (1995) and den Dikken (2005) as a root clause which serves as the comment of the above-mentioned topic, and which provides information about the focal element. This account assigns to (3a) the structure in (6) (cf. den Dikken 2005: 698). S1 in (6) is the topic clause harboring the focus expression, something, and S2 is the root clause making a comment on the focus expression in S1.

5 Certain English dialects, however, do allow SCRs in these positions; see Doherty (2000: 87-89) on these varieties. See also Henry (1995: 125-126) for the distribution of SCRs in Belfast English.

6 Note, however, that in den Dikken’s (2005) analysis, this focus is not represented as occupying a syntactic position.

7 Note, however, that according to den Dikken (2005: 700), not every SCR can be analyzed...
Syntactically, den Dikken (2005: 698, fn. 6) compares $S_1$ to hanging topics, whose function is also to introduce (the referent of) a nominal into the discourse. Hanging topics are not syntactically integrated in the root clause which provides the comment. Observe that the matrix in the comment clause can be omitted (compare (7a) with (7b), cf. Henry 1995: 132, Haegeman 2015), rendering hanging topic constructions structurally akin to SCRs.

(7)  
   a. As for syntax, I found it too difficult.  
   b. As for syntax, $\emptyset$ found it too difficult.

The reader is referred to Henry (1995: 131-135) and den Dikken (2005) for further details of this line of analysis. The relevant point to retain is that the modifying clause ($S_2$) does not form a constituent with the nominal head (or the ‘focus expression’, according to den Dikken 2005), to the exclusion of the material which occurs linearly to the left of the nominal expression, as shown in (6).

Recently, Haegeman (2015; see also Haegeman et al. 2015) has evaluated the topic-comment representation of SCRs. By adducing a number of observations on their internal and external syntax, she concludes that a topic-comment analysis of SCRs cannot be maintained. In what follows, I will list two of her observations on the internal syntax of SCRs, which are relevant for the initial analysis of $tu$-less subject RCs in PhG.

First, Haegeman (2015) observes that, were a SCR to instantiate a topic-comment structure in which ($S_2$) and the nominal head (focus) do not form a constituent, then the head and the modifying clause would not be expected to enter in a coordinate relation with a regular RC. However, she shows that this prediction is not borne out: examples such as (8) are judged grammatical. This provides evidence for Doherty’s (1994, 2000) analysis that the nominal head and

this way.
the RC form one structural unit.

(8) I have [one colleague runs a sushi shop] and [another one who has a burger restaurant].

(Haegeman 2015: 140, her (16a))

Second, she observes that, according to the configuration in (6), any constituent contained within $S_1$ is not expected to c-command (any material contained in) $S_2$; however, this is not necessarily the case. In (9), for instance, the pronoun his, contained within the SCR, receives a bound-variable reading, which is possible only if it is in the c-commanding domain of the quantified subject, every student, in $S_1$.

(9) Every student, is looking for a teacher speaks his, language.

(Haegeman 2015: 141, her (17a))

The availability of the bound-variable reading of the pronoun in (9) provides additional support for the RC analysis of SCRs, because this approach correctly predicts that the RC is within the c-command domain of the matrix subject.

Based on the above observations (among several others), Haegeman (2015) concludes that a topic-comment analysis of SCRs in English is not tenable, and that her observations offer support for any argument which recognizes SCRs as (some type of) RC.

In the next section, I return to PhG subject RCs without a complementizer.

3 Are there SRCs in PhG?

Unlike English, PhG is a consistent null-subject language: there are no expletives in existential clauses with a postverbal subject for instance (10a); similarly, weather verbs never have overt subjects (10b). As expected, referential pronominal subjects can also be omitted (10c).

(10) a. Ísanti dhíu néčis sin stráta.  
were.3PL two women on-the road  
‘There were two women on the road.’

b. (*Ató) vrešízi. 
It rain.3SG  
‘It is raining.’

c. (Até) kathéti ačí.  
She sit.3SG there  
‘She is sitting there.’
Given that pronominal subjects can freely be omitted (10c), it is at first blush difficult to conclude whether (2) is a genuine RC, as Anastasiadis (1976) suggests, or whether it involves two independent sentences with null subjects, as in (11).

(11) Ø ídha an goritsí. Ø kathéti ačí monaxó ts.
saw.1sg a girl sit.3sg there on-her-own
‘I saw a girl. She is sitting there alone.’ (cf. (2))

Informants agree with Anastasiadis (1976) and accept (2) with a RC reading: they state that there is no intonation break between goritsí ‘girl’ and kathéti ‘(she) is sitting’, contrary to the case in (11). Furthermore, for these informants, (12a) and (12b) are associated with distinct readings: (12a) where two sentences are simply placed side by side, means there were two, and only two, women in the car and they were both wearing necklaces. On the other hand, (12b), under a RC reading, is reported to imply the existence of other women in the car as well, who were, however, not wearing necklaces.

(12) a. So tomofíli ísanti mo dhíu néčis. Forénkanti kerdanníxi.
in-the car were.3pl only two women wore.3pl necklace.
‘In the car, there were only two women. They were wearing necklaces.’

b. So tomofíli ísanti mo dhíu néčis forénkanti kerdanníxi.
in-the car were.3pl only two women wore.3pl necklace.
‘In the car, there were only two women who were wearing necklaces.’

Structurally, the reading in (12a) suggests that the only focalized constituent c-commanded by the focus sensitive exhaustivity operator mo ‘only’ is the DP dhíu néčis ‘two women’; the second sentence forénkanti kerdanníxi ‘they were wearing necklaces’, however, is not in the c-commanding domain of mo ‘only’ as operator-focus dependencies do not cross sentence boundaries. As a result, in the first sentence, the number of the women that were in the car is strictly defined as “2”. The second sentence then merely provides additional information on the only two women that were in the car. The reading in (12b), on the other hand, suggests that the constituent forénkanti kerdanníxi ‘were wearing necklaces’ functions as a restrictor of the set of referents for the DP dhíu néčis ‘two

---

8 See further Andriotis (1948: 51) and Anastasiadis (1976: 248), who do not put any punctuation mark (comma, period etc.) between the head noun and the associated modifying clause in examples of SCRs, which would possibly mark an intonation break.
women’; in turn, the string *dhíu néčis forénkanti kerdanníxi* ‘two women (who) were wearing necklaces’ is c-commanded by *mo* ‘only’ as one single constituent. Due to the existence of the restricting clause, the sentence receives a reading in which the number of women that were at the same time both in the car and wearing necklaces was only “2”. This implies that other women might have also been in the car, who were not wearing necklaces. Based on the evidence (12b) provides, we may preliminarily conclude that *tu*-less restrictive subject RCs exist in PhG, as Anastasiadis (1976) suggests.

The omission of *tu* ‘that’ in subject RCs is not entirely free, however. The environments which can harbor *tu*-less subject RCs are near-identical to those in which English SCR are tolerated: *tu* ‘that’ omission is most naturally tolerated in complements of the predicate *thóró* ‘see’, when this predicate introduces a new referent into the discourse (2), in copular existential sentences (12b), and in *have* existential sentences (13a). Preliminary judgments suggest that, in these cases, *tu* ‘that’ can be deleted as long as the antecedent is indefinite, cf. (13a) with (13b).

\[
\begin{align*}
(13) & \quad \text{a. } \text{Éxu} & [a \text{ jos (tu) kamnóni sa } \text{Ádhana}]. \\
& \text{have.1SG a son that work.3SG in-the Adana} \\
& \text{‘I have a son who works in Adana’}. \\
& \text{b. } \text{Dáma mu } \text{íxa} \ [\text{to jo } \text{mu *(tu) kamnóni sa } \text{Ádhana}]. \\
& \text{with me had.1SG the son my that work.3SG in-the Adana} \\
& \text{‘With me, I had my son who works in Adana’}.
\end{align*}
\]

Speaker judgments reveal that subject RCs without the complementizer *tu* ‘that’ are severely degraded when in subject or indirect object positions:

\[
\begin{align*}
(14) & \quad \text{a. } \text{[An gorítsi *(tu) kathéti ačí monaxó ts] irévi xalxás.} \\
& \text{a girl that sit.3SG there on-her-own want.3SG bagel} \\
& \text{‘A girl who is sitting there on her own wants a bagel.’}
\end{align*}
\]

\[
\begin{align*}
(13) & \quad \text{b. } \text{[An gorítsi *(tu) kathéti ačí monaxó ts] irévi xalxás.} \\
& \text{a girl that sit.3SG there on-her-own want.3SG bagel} \\
& \text{‘A girl who is sitting there on her own wants a bagel.’}
\end{align*}
\]

9 It should be noted that *tu* ‘that’ omission is allowed only when the RC is a postnominal one, cf. (2) with (i):

\[
\begin{align*}
(13) & \quad \text{a. } \text{[An gorítsi *(tu) kathéti ačí monaxó ts] irévi xalxás.} \\
& \text{a girl that sit.3SG there on-her-own want.3SG bagel} \\
& \text{‘A girl who is sitting there on her own wants a bagel.’}
\end{align*}
\]

10 The examples cited in Andriotis (1948: 51) and Anastasiadis (1976: 248) also verify these generalizations.
b. Pítaksa an xalxás s [an gorítsi *(tu) kathéti ačí monaxó ts].
   sent.1SG a bagel to a girl that sit.3SG there on-her-own
   ‘I sent a bagel to a girl who was sitting there on her own.’

When the distributional similarity between English SCRs and PhG RCs that allow complementizer omission is taken into consideration, we can tentatively conclude that a structure such as (2) can be referred to as a SCR.

There are three (additional) pieces of evidence that suggest that in PhG a SCR forms a constituent with its associated nominal ‘head’, which, in turn, suggest that SRCs are structurally embedded, rather than independent sentences.

First, a SRC can (marginally) occur as a left-peripheral topic, when the predicate of the matrix clause is thoró ‘see’:

(15) ?[An av néka (tu) ifanéni an t esénaj], čo ídha tač so
   an other woman that weave.3SG like you not saw.1SG her in.the
   ómbri mu.
   life my
   ‘Another woman who weaves like you, I have not seen in my life.’

Second, similar to Haegeman’s (2015) observation for English SCRs, a SCR in PhG can also be coordinated with a regular relative clause, as shown in (16).

(16) Éxu [a jos (tu) kamnóni sa Ádhana] če [an góri tu
   have.1SG a son that work.3SG in-the Adana and a daughter that
   paredhóthin so Everéki].
   married.3SG in-the Everek
   ‘I have a son who works in Adana and a daughter who is married in Ev-
   erek.’

Finally, again similar to the case of English SCRs as observed by Haegeman (2015), a quantificational element c-commanding the (head of) the SRC can bind a pronominal element inside the RC, giving rise to a distributive interpretation for the pronoun:

(17) [Xer o mástrus], éši [a čiráxus (tu) katéši ti
   every the master have.3SG an apprentice that understand.3SG the

---

The observations that follow also argue against a possible topic-comment analysis of SCRs in PhG. Nevertheless, as I have not investigated all the possible predictions of such an analysis in PhG, I do not want to make reference to it in the rest of this note.
gwósa tu.
language his
‘Every master has an apprentice who understands his language.’

Under the reading in which (37) involves two independent sentences, the bound variable interpretation of the pronoun tu ‘his’ is no longer available, suggesting that the pronoun is not c-commanded by the quantified subject of the first sentence:

(18) [Xer o mástrus], éši a čiráxus. Katéši ti
every the master have.3SG an apprentice understand.3SG the
gwósa tu\*,\[i/=j,
language his
‘Every master has an apprentice. He understands his language.’

4 Conclusions
In this brief note, I have provided a preliminary analysis of subject RCs in PhG which allow for complementizer deletion. Based on distributional evidence, I concluded that the PhG structures under discussion are very similar to English SCR s. In relation to the internal syntax of PhG complementizerless RCs, I concluded that the relevant structures form a constituent with their nominal head, much as is the case in canonical (head-external) RCs. The precise structure of complementizerless RCs in PhG, the exact nature of the environments they are tolerated in, and how they differ from RCs with an overt complementizer are issues that await further research.

References
Anastasiadis, V. 1976. I sintaksi sto Pherasiotiko idioma tis Kappadokias [The syntax of the dialect of Pharasa in Cappadocia]: University of Ioannina dissertation.
Subjunctive selection in French vs. Balkan

Lena Baunaz & Eric Lander

1 Verbal and clausal mood

Languages vary as to how and where they mark subjunctive vs. indicative mood (Sočanač 2017, among others). Whereas French marks the indicative vs. subjunctive mood on the verb (1), Balkan languages use a special (so-called) complementizer to indicate subjunctive mood (2) (see Giannakidou 1998, 2009 and subseq. as well as Roussou 2000, 2009, 2010 and subseq. for Modern Greek; see Sočanač 2017 for details on Slavic and Balkan languages).

(1) a. Mirka dit que Roger est prêt à l'heure. (French) M. says that R. is ready on time
b. Mirka ordonne que Roger soit prêt à l'heure. M. orders that R. be.SUBJ ready on time

(2) a. Nomizo oti kerdizei o Janis. (MG) think1SG that win3SG.IMPERF the John
   ‘I think that John is winning.’
b. Thelo na kerdisi o Janis want1SG that.SUBJ win3SG.PERF the John
   ‘I want John to win.’ (Giannakidou 2009: 1887)

Most of the time, indicative and subjunctive mood morphology are syncretic with one another in French, but certain verbs—such as être ‘be’—retain distinct subjunctive morphology. In Modern Greek (MG) and Balkan more generally,
subjunctive morphology has been lost altogether and been replaced by what is known as *perfective non-past* (PNP) morphology, which encodes tense and aspect but not mood (mood being marked on the so-called complementizer).

The French strategy of marking subjunctive on the embedded verb is traditionally called *verbal mood*; the Balkan strategy of marking subjunctive on the complementizer may be referred to as *clausal mood* (see Sočanać 2017 and references there).

### 2 Veridicality and Mood selection

It is generally accepted that the subjunctive mood is a *dependent mood* selected under verbs which are associated with some special semantic features (Quer 1998, 2001, 2009, Giannakidou 1998, 2009, among others). Giannakidou, using MG data, argues that veridicality, as defined in (3), licenses mood choice:

(3) **Veridicality** (Giannakidou 1998, 2009)

A propositional operator $F$ is veridical iff from the truth of $Fp$ we can infer that $p$ is true according to some individual $x$ (i.e. in some individual $x$’s epistemic model).

According to (3), an embedded proposition has to be true for at least one individual (the subject of the main verb and/or the speaker), in all the worlds of a relevant model. Giannakidou claims that veridicality triggers the indicative mood in the embedded clause and non-veridicality triggers the subjunctive mood.

The correlation between mood and veridicality applies very well in MG and most of the Balkan Slavic languages (see Todorovic 2012), but it fails to apply to Romance’s emotive factive complements (Quer 1998, 2001, 2009, Baunaz & Puskás 2014, Baunaz 2015, 2017). In Romance, here illustrated with French, predicates like *regretter* ‘regret’, *être content* ‘be happy’, etc. unexpectedly trigger the subjunctive mood, as seen in (4), even though these predicates are veridical according to (3).

---

There is a long-standing debate on the status of *na* in the syntactic literature: is it a mood particle (i.e. an inflectional mood head, see Philippaki-Warburton 1994, 1998, Giannakidou 1998, 2009 among many others), a complementizer (Agourakis 1991, Tsoulas 1993 and Roussou 2000, a.o.) or a hybrid, that is a mood head moving to C (Giannakidou 2009)? See Giannakidou & Mari (2017) for a state of the art on this topic. See below for our take on the matter.
Mirka regrette que Roger ait perdu. (French) Mirka regrets that Roger has lost.

It appears that Giannakidou’s definition of veridicality cannot account for mood distribution cross-linguistically. Baunaz & Puskás (2014) refine (3) by investigating the notion of “some individual”, as applied to six verb classes of embedding verbs in French. Baunaz (2018) extends their findings to Balkan languages. They argue that these predicates can be classified into three groups if one takes into account the idea that the truth of the embedded proposition can relate to both the subject and the speaker (= strong veridicality), to either the subject or the speaker (= relative veridicality), or to none of them (= non-veridicality). Note that verbs can be “ambiguous” in being either strongly veridical or relative-veridical (MG thimame ‘regret’ or French comprendre ‘understand’ are cases in point; see Table 1).

<table>
<thead>
<tr>
<th>English tr.</th>
<th>MG</th>
<th>SC</th>
<th>Bg</th>
<th>Fr</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Strong veridical</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘remember’</td>
<td>thimame</td>
<td>sjetiti se</td>
<td>pomnja</td>
<td>se rappeler</td>
</tr>
<tr>
<td>‘regret’</td>
<td>thimame</td>
<td>žaliti</td>
<td>sǎžaljavam</td>
<td></td>
</tr>
<tr>
<td>‘understand’</td>
<td>thimame</td>
<td>sjetiti se</td>
<td>pomnja</td>
<td>comprendre</td>
</tr>
<tr>
<td><em>Relative veridical</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘remember’</td>
<td>lipame</td>
<td>žaliti</td>
<td>sǎžaljavam</td>
<td>regretter</td>
</tr>
<tr>
<td>‘regret’</td>
<td>lipame</td>
<td>žaliti</td>
<td>sǎžaljavam</td>
<td>comprendre</td>
</tr>
<tr>
<td>‘understand’</td>
<td>lipame</td>
<td>žaliti</td>
<td>sǎžaljavam</td>
<td>comprendre</td>
</tr>
<tr>
<td><em>Non-veridical</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘say’</td>
<td>leo</td>
<td>reći</td>
<td>kazvam</td>
<td>dire</td>
</tr>
<tr>
<td>‘want’</td>
<td>thelo</td>
<td>željeti</td>
<td>iskam</td>
<td>vouloir</td>
</tr>
</tbody>
</table>

Table 1: Some (non-)veridical verbs in Modern Greek, Serbo-Croatian, Bulgarian and French (from Baunaz 2018)

The distinction between strong and relative veridicality is tracked by different complementizers in Balkan (see Table 3), or by subjunctive vs. indicative mood on the embedded verb in French (see Table 2). The three-way distinction above does not provide an orderly way of accounting for mood selection in French, though: strong-veridical verbs and some non-veridical verbs select for indicative complements, whereas relative-veridical and some non-veridical verbs select for subjunctive complements. This results in an *ABA violation, as indicated by shading in Table 2. In MG (and Balkan more generally), veridicality and mood-marking appear to be unrelated too: non-veridical verbs can take an oti-clause with an indicative embedded verb, or they can ap-
pear with subjunctive *na*-complements. This is summed up in Table 2:

<table>
<thead>
<tr>
<th>Main Predicates</th>
<th>Complementizer</th>
<th>Mood on V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fr</td>
<td>MG</td>
</tr>
<tr>
<td>Strongly veridical</td>
<td>que</td>
<td>pu</td>
</tr>
<tr>
<td>Relative veridical</td>
<td>que</td>
<td>pu/oti</td>
</tr>
<tr>
<td>Non-veridical 1</td>
<td>que</td>
<td>oti</td>
</tr>
<tr>
<td>Non-veridical 2</td>
<td>que</td>
<td>na subjunctive</td>
</tr>
</tbody>
</table>

Table 2: Mood and veridicality in French and Modern Greek

The tripartition does, however, show promise in terms of complementizer selection, especially by splitting the non-veridical group in two (an ‘indicative’ NV1 and a ‘subjunctive’ NV2). Table 3 illustrates complementizer selection under the relevant predicates in the four languages at stake.

<table>
<thead>
<tr>
<th></th>
<th>MG</th>
<th>Bg</th>
<th>Serbian</th>
<th>Croatian</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly veridical</td>
<td>pu</td>
<td>deto</td>
<td>što</td>
<td>da</td>
<td>que</td>
</tr>
<tr>
<td>Relative veridical</td>
<td>pu</td>
<td>če</td>
<td>što</td>
<td>da</td>
<td>que</td>
</tr>
<tr>
<td>Non-veridical 1</td>
<td>oti</td>
<td>če</td>
<td>da</td>
<td>da</td>
<td>que</td>
</tr>
<tr>
<td>Non-veridical 2</td>
<td>na</td>
<td>da</td>
<td>da</td>
<td>da</td>
<td>que</td>
</tr>
</tbody>
</table>

Table 3: Finite complementizers in Modern Greek, Bulgarian, Serbian, Croatian, French

There are no problematic *ABA violations in Table 3. In other words, veridicality cleanly tracks complementizer selection in French and Balkan.

Focusing on Balkan languages, Baunaz (2015, 2016, 2018) argues that the complementizer morpheme has an internal functional sequence (fseq), and that the complementizer may lexicalize differently sized structures from this fseq. She also claims that the distribution of complementizers is governed by veridicality and shows that veridicality plays a role in so-called factive islands, called *veridicality islands* in Baunaz (2018). Based on syncretism patterns with complementizers and on wh-extractions out of (non-)veridical domains in four different languages, she argues that the predicates in Table 1 select for complementizers of different ‘sizes’. She claims that the size of the complementizer plays a role in strong, weak, or non-island configurations. The basic complementizer-fseq...
resulting from Baunaz’s work is illustrated in (5). Note that under this analysis, na (and its cognates in Balkan languages) is considered to be a complementizer of the smallest size, with distinct syntactic properties from, say, non-veridical complementizers (see fn. 1).

(5) strong-veridical > relative-veridical > non-veridical 1 > non-veridical 2

Baunaz also adopts the idea from Manzini & Savoia (2011) and Roussou (2010) that complementizers may introduce a propositional operator (OP_{prop}). This operator closes off the clause and turns it into a proposition. Subjunctive clauses, which are open propositions, do not involve such an operator.

In this paper, we would like to reconcile the theory of complementizer selection summed up in Table 3 with the way in which mood is realized (Table 2). We will explain the apparent *ABA in French as well as some of the properties of the MG mood complementizer na (and its Balkan cognates, as described in Soćanač 2017) by taking a nanosyntactic approach.

3 The analysis: Peeling and packaging

In our view, the apparent ABA in French, seen in Table 2, is only an illusion. It is the result of unduly mixing up two independent processes: on the one hand, the internal structure of complementizers in terms of veridicality, and on the other hand, the selection of embedded mood by certain matrix verbs. Taken on its own, neither process violates the *ABA theorem. As we saw above in Table 3, the veridicality domain does not show any complementizers in an illicit ABA-type configuration. And in Table 4 we see that mood is triggered by the kind of selecting predicate, i.e. the internal structure of the predicate.

As seen in Table 4, indicative and subjunctive obey the adjacency requirement on syncretism, that is, they do not show any illicit ABAs.

The problem arises in the observation that a relative-veridical complementizer appears to ‘select’ a subjunctive embedded verb. However, the correct way of framing this fact is not that the complementizer selects the embedded mood; rather, it is the higher matrix verb that is responsible for this. Veridicality is an (at least partially) independent variable, with different complementizer structures being compatible with different matrix verbs (6).

---

2 For reasons of space, we borrow Baunaz & Puskás’s (2014, under review) classification of embedding verbs without explanation. The reader is referred to their work for details.
Table 4: Matrix predicate classes and embedded mood

<table>
<thead>
<tr>
<th>Predicate class</th>
<th>Internal structure</th>
<th>Selects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directives</td>
<td>CAUSE &gt; VOLITIONAL &gt; EMOTIVE &gt; SENTIENT</td>
<td>SUBJ</td>
</tr>
<tr>
<td>Desideratives</td>
<td>VOLITIONAL &gt; EMOTIVE &gt; SENTIENT</td>
<td>SUBJ</td>
</tr>
<tr>
<td>Emotive factives</td>
<td>EMOTIVE &gt; SENTIENT</td>
<td>SUBJ</td>
</tr>
<tr>
<td>V. of saying/epistemic</td>
<td>SENTIENT</td>
<td>IND</td>
</tr>
</tbody>
</table>

(6) Directives    (→ SUBJ) SV-Comp
Desideratives    (→ SUBJ) RV-Comp
Emotive factives (→ SUBJ) NV₁-Comp
Saying/epistemic (→ IND) NV₂-Comp

As seen in (6), RV and NV₂ complementizers are compatible with verbs taking the subjunctive mood (directives, desideratives and emotive factive verbs); SV and NV₁ are compatible with verbs taking the indicative (verbs of saying and epistemic verbs).

We make two main assumptions in our analysis. First, selection is local. Second, the order in which the main elements of the biclausal configuration are generated is not the following:

(7) Conventional ordering (not adopted here):
   i. [embedded verb/clause]
   ii. [Comp [embedded verb/clause]]
   iii. [matrix verb/clause [Comp [embedded verb/clause]]]

Instead we will assume that veridicality – the stance taken by some individual with regard to a proposition – is in some sense the ‘core’ of the sentence; therefore we will hypothesize that the complementizer, as encoder of veridicality, is generated first.

(8) [RV [NV₁ [NV₂ [C]]]] (relative-veridical Comp)

In (8) we have provided the relative-veridical structure, but any complementizer
structure would also be available: strong-veridical [SV [[RV _ [NV₁ _ [NV₂ _ [C]]]]]], or non-veridical [NV₁ _ [NV₂ _ [C]]] or [NV₂ _ [C]]. We have chosen the relative-veridical structure because it partakes in the problematic ABA pattern discussed above.

The next step, for our purposes, is external merge of the verb with inflectional layers above (TAM, where \textit{Subj} stands for subjunctive mood; \textit{Prop} stands for proposition, which is also the feature yielding indicative morphology). This is seen in (9).

(9) \[ \text{Prop [Subj [T [Asp [V]]]] + [RV [NV₁ _ [NV₂ _ [C]]]]} \]

Following a ‘peeling’ approach to selection (unpublished work by Michal Starke, Caha 2009: ch. 4; see Taraldsen Medová & Wiland 2018 for a recent implementation), we will take the full structure of the verb to be merged at this stage, with selection of a subset of this structure happening in a later step, leaving behind a so-called peel.

In the next stage of the derivation the matrix verb is merged. As discussed above, it is the class of the matrix verb that determines whether the embedded mood is indicative or subjunctive. In this case, with the relative-veridical Comp, we have an emotive-factive matrix predicate.

(10) \[ [V\text{\textsubscript{matrixP}}] + [\text{Prop [Subj [T [Asp [V]]]]} + [RV [NV₁ _ [NV₂ _ [C]]]]} \]

here: \( [V\text{\textsubscript{matrixP}}] = [\text{EMOTIVE [SENTIENT [VP]]}] \)

At this stage of the derivation, the matrix VP is in a local relation to the embedded verb; we posit that this kind of relation makes it possible for the matrix verb to select some subset of the verbal structure, which here would be a verb with subjunctive inflection.

Selection involves movement of the subset to the left of the matrix VP, as sketched in (11).

(11) \[ [\text{Subj [T [Asp [V]]]}] + [V\text{\textsubscript{matrixP}}] + [\text{Prop [ __ ]}] + [RV [NV₁ _ [NV₂ _ [C]]]] \]

Importantly, removing the subjunctive structure leaves behind a Prop layer without its complement (since SubjP has been moved out); this Prop ‘peel’ can be spelled out as part of the complementizer itself, i.e. Fr. \textit{que}, as seen here:

\footnote{3 See Baunaz (2017) and Baunaz & Puskás (under review) for arguments in favor of the idea that what determines the subjunctive mood is the emotive feature in the fseq of the matrix verb.}
This means that the structure of French *que* as it is stored in the lexicon would look like (13).
In other words, the Prop peel encodes movement of a subjunctive verb out from underneath. However, it is not only in such cases that que can be spelled out. *Que* can also be spelled out if an indicative verb has been selected/moved to the left of the matrix VP, as seen in (14).
The lexical entry in (13) can still, by the Superset Principle, spell out the leftover structure in (14) as que. That is to say, the lexical structure with the peel in (13) is still a superset of [SV [RV [NV₁ [NV₂ [C]]]]] (or any subset of this structure). In other words, que is spelled out whether indicative or subjunctive has been selected by the matrix verb.

Thus the matrix verb does not directly determine the kind (size) of the complementizer. In other words, the complementizer grows and shrinks independently of the behavior of the matrix verb and the process of mood-selection happening above it. The lexical entry in (13) is able to accommodate this fact if we adopt the Revised Superset Principle, given in (15).

---

4 We assume that later reshuffling in the derivation will move [V_{matrix} P [Comp]] to the left of the embedded verb/clause, giving us the order [matrix - Comp] embedded ____. Furthermore, within the V-zone, we assume that VP eventually moves up to the left of the layers above it, giving the expected order [verb] - INFL ____.
A lexical entry L may spell out a syntactic node SN iff the features of L are a superset of the features dominated by SN.

This relaxation of the ‘classical’ Superset Principle is required on independent grounds (see Vanden Wyngaerd 2018 for more details). For our purposes here, (15) allows for the lexical structure of que in (13) to map, at least in principle, onto all of the following syntactic structures:

(16) a. subjunctive verb selected

```
[ [Prop __ ] [SV [RV [NV1 [NV2 [C]]]]] ]
[ [Prop __ ] [RV [NV1 [NV2 [C]]]] ]
[ [Prop __ ] [NV1 [NV2 [C]]] ]
[ [Prop __ ] [NV1 [NV2 [C]]] ]
```

“shrinking in the middle”

b. indicative verb selected

```
[SV [RV [NV1 [NV2 [C]]]]]
[RV [NV1 [NV2 [C]]]]
[NV1 [NV2 [C]]]
[NV2 [C]]
```

In (16a) we indicate that the complementizer zone can shrink while still keeping the Prop peel in place at the top; this is a kind of “shrinking in the middle” (as Vanden Wyngaerd 2018 puts it), a feature of the Revised Superset Principle which is absent in the traditional Superset Principle.

Turning now to Greek, it turns out that our analysis of French has some interesting consequences for the subjunctive non-veridical (= [NV₂ [C]]) complementizer-like particle MG na. First of all, these kinds of complements are not propositional and thus not expected to build all the way up to Prop, rather the structure stops at Subj. Second, recall that the embedded verb does not display a morphological subjunctive but rather perfective non-past inflection. For our purposes, PNP corresponds to [Asp [V]].

45
Crucially, once this part has been extracted from the verbal structure, we are left with a peel made up of Subj and T. Thus the lexical entry for *na* is the one given in (18).

(18) \[ \langle na \leftrightarrow [\text{Subj} [T \_]] [\text{NV}_2 [C]] \rangle \]

This entry accounts for two major properties of *na*, namely that it is usually taken as a marker of subjunctive and also that it is an infinitival marker in T – a syncretism which is in fact widespread in Balkan.

Zeroing in on the C-zone, note that the lexical structure of *na* contains only [\text{NV}_2 [C]] and not a larger set of veridicality features (as in the lexical structure of Fr. *que*). The higher veridicality features are instead spelled out in MG by *pu* and *oti*. Embedded verbs under *pu* and *oti* cannot take PNP but instead show indicative morphology. Furthermore, contrary to *na*-clauses, *pu*- and *oti*-clauses
are propositional. Thus we can assume for pu and oti that the full PropP structure has been extracted in these contexts, leaving no peels behind. The lexical entries we land on are:

(19)

a. \langle pu \leftrightarrow [SV [RV [NV_{1} [NV_{2} [C]]]]] \rangle \\
b. \langle oti \leftrightarrow [RV [NV_{1} [NV_{2} [C]]]] \rangle \\

The lack of peels in (19), then, encodes the fact that a larger verbal structure (namely indicative, requiring Prop in our analysis) occurs under pu and oti. The bigger peel in the lexical structure of na, on the other hand, encodes the fact that a smaller verbal structure (PNP, basically just a species of aspectual marking) occurs under na.

4 Conclusions

Nanosyntax makes it possible to ‘package’ features (and then store them in the lexicon) in different ways cross-linguistically. For us, it is crucial that Subj can be spelled out on the embedded verb in French, which has verbal mood, but as part of the complementizer na in MG, where na is a manifestation of clausal mood. This general nanosyntactic strategy has in this paper been combined with a peeling approach to selection, whereby a leftover peel from the verbal zone implies what kind of inflected verb has been extracted. We show that packaging and peeling, combined with the assumption that the complementizer is built before the embedded verb/clause, can be utilized to coherently account for the facts of subjunctive selection in French vs. Balkan.

References

Baunaz, L. 2017. French predicates selecting the subjunctive mood under the microscope: The emotive factor. In S. Perpiñán, D. Heap, I. Moreno-Villamar


‘U is’ or ‘U bent’?
Polite forms of address and verbal agreement in historical Dutch and Ibero-Romance*

Miriam Bouzouita & Anne Breitbarth

1 Introduction

It is well-known that the (polite) forms of address in Dutch underwent a cyclic replacement by which the 2nd person nominative plural pronoun *ghi/gij* became the formal 2nd nominative singular pronoun (V(os)-form), and later the informal 2nd nominative singular pronoun (T(u)-form), replacing older *du*, while the 2nd person accusative plural pronoun *u* first became the accusative of the new singular V-form, and later the nominative V-form. The developments are summarised in Table 1 (adapted from [van Leuvensteijn 2002: 289 and Vermaas 2005]).

Table 1: Cyclic shifts in 2nd person pronouns in the history of Dutch

<table>
<thead>
<tr>
<th></th>
<th>Sg</th>
<th>Pl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nom</td>
<td>Acc/Dat</td>
</tr>
<tr>
<td>Middle Dutch (c. 1150-1500)</td>
<td>du</td>
<td>di</td>
</tr>
<tr>
<td>1500-1700</td>
<td>du, ghi</td>
<td>di, u</td>
</tr>
<tr>
<td>17th/18th c.</td>
<td>gij</td>
<td>u</td>
</tr>
<tr>
<td>today</td>
<td>V: u</td>
<td>&lt; gij lliend</td>
</tr>
<tr>
<td>T: je/jij</td>
<td>je/jou</td>
<td>jullie</td>
</tr>
</tbody>
</table>

*We acknowledge invaluable input from our fellow ΔiaLing colleague Jacques Van Keymeulen to earlier versions of this paper.
A question that has not received much attention at all in the literature is how the verbal agreement with these pronouns evolved. There are two theories regarding the rise of nominative u that make different predictions with respect to this question. First, according to e.g. Vor der Hake (1911), the nominative use of u arose from the accusative one by semantic shift. Second, as proposed by Van der Horst (2008: 1094), the nominative use of u arose via an intermediate step, namely the epistolary forms of address consisting of a (possessive) pronoun and a honorific noun, U.Ed., U.E. (uw edelheid, ‘your honour’) and U.L. (uwe(r) liefde ‘your love / kindness’) used in letters from the 17th century onwards. These were first only used in writing, originally in chancery style for nobility, and then oralised as uwé [uw‘eː] / [‘uwε], and spread top-town through social classes (Kern 1911, Heeroma 1934). The first theory predicts that 2nd person agreement on the verb with subject-u should be older, as the accusative form of the 2nd person pronoun, when it is reanalysed as nominative, is still a 2nd person pronoun. Under the second theory, one would expect that u(wé) should first have occurred with 3rd person agreement, as the form of address is a noun phrase. Van der Horst bases this latter theory on a comparison with a similar development in German, where the new V-form Sie ‘they > you(V)’, argued to be a pronominalization of IhrGnaden ‘your mercies’ (e.g. Simon 2003), and goes with 3rd person agreement on the verb.

Our paper departs from three observations. First, there is variation in present-day Dutch regarding the agreement morphology on the finite verb that goes with the V-form u: where this morphology is not syncretic (as it is in most verbs), it varies between 2nd and 3rd person agreement. In (1), this is illustrated for the verbs hebben ‘have’, zullen ‘shall’, and kunnen ‘can’.

(1) 2nd person: u hebt/zult/kunt
3rd person: u heeft/zal/kan
‘you (V) have/shall/can’

Second, the earliest (pre-1600!) occurrences of u used as a nominative pronoun that are reported in the literature occur with 2nd person agreement on the verb (Paardekooper 1996). This is a problem for Van der Horst’s theory of the develop-

---

1”Als een briefschrijver de geadresseerde met Uwe Edelheid aanspreekt, gebruikt hij een substantiegroep met Edelheid als kern. Hier horen de persoonlijke, bezittelijke en wederkerende voornaamwoorden van de derde persoon bij.” (van Leuvensteijn 2002: 290)

2We are not considering the variation in colloquial Belgian Dutch regarding the use of u in the present paper.
opment of the verbal agreement with u.\footnote{Examples quoted after Paardekooper (1996: 70).}

(2) *ende u hebt in dese wech al 7 vame ende bij lant komende 5 vamen*  
And you have in this way already 7 fathom, 6 fathom, and, coming onto the shore, 5 fathom.\footnote{[1599] De tweede schipvaart der Nederlanders naar Oost-Indië [...]}

(3) *Wat belangt de regerijng is off soude te lanck sijn te verhaelen, dat sult v alles verstaen godt ons met lijff in Nerlant helpende*  
Concerning the government, it is or would be too long to tell (all), you shall understand all of that, if God help us with life in the Netherlands.\footnote{[1599] B. Journaal van Reijer Cornelisz.}

The third observation comes from outside the Germanic sphere. To be more precise, Lara Bermejo\footnote{Lara Bermejo (2015, 2016a, b)} observes that in some Ibero-Romance (IR) varieties, such as Western Andalusian Spanish, the plural V-form *ustedes* has replaced the plural T-form *vosotros*, but the agreement of e.g. the reflexive pronouns, object clitics (os > se) and the finite verbs lag behind, varying between 2\textsuperscript{nd} and 3\textsuperscript{rd} person agreement, as illustrated in (4). He also sketches a geographical diffusion pattern, whereby the centre of innovation is situated in the Cádiz province, which has more of the innovative 3\textsuperscript{PL} features closer to the center (level 4) than those areas that are further away from it (levels 1 (furthest away, least advanced) to 3).

<table>
<thead>
<tr>
<th>Sg</th>
<th>Pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modern Peninsular Spanish</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>V</td>
</tr>
<tr>
<td>V</td>
<td>Ustedes</td>
</tr>
<tr>
<td>Modern Western Andalusian</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>&gt;</td>
</tr>
<tr>
<td>V</td>
<td>Usted</td>
</tr>
</tbody>
</table>

(4) *Ustedes sois hermanos.*  
‘You(3\textsuperscript{PL}) are(2\textsuperscript{PL}) siblings’ \footnote{Lara Bermejo 2016b: 98}

As can be seen from Table 2, Lara Bermejo identifies a cline following which the *ustedes* phenomenon progresses from one syntactic context to another, whereby
Table 2: Extension of the innovative 3PL in the ustedes phenomenon (Lara Bermejo 2016a: 277)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Stressed pronoun</th>
<th>Reflexive pronoun</th>
<th>Verbs</th>
<th>Accusative pronoun</th>
<th>Dative pronoun</th>
<th>Possessives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>3pl</td>
<td>2pl</td>
<td>2pl</td>
<td>2pl</td>
<td>2pl</td>
<td>2pl</td>
</tr>
<tr>
<td>Stage 2</td>
<td>3pl</td>
<td>3pl</td>
<td>2pl</td>
<td>2pl</td>
<td>2pl</td>
<td>2pl</td>
</tr>
<tr>
<td>Stage 3</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>2pl</td>
<td>2pl</td>
<td>2pl</td>
</tr>
<tr>
<td>Stage 4</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>2pl</td>
<td>2pl</td>
</tr>
<tr>
<td>Stage 5</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>2pl</td>
</tr>
<tr>
<td>Stage 6</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
<td>3pl</td>
</tr>
</tbody>
</table>

Stage 6 sees full completion of 3PL, which manifests itself in Canarian and Latin American Spanish.

In the present paper, we report on a pilot study on the historical development of the verbal agreement going with the incipient use of epistolary forms of address and emerging nominative u in a corpus of letters (Letters as Loot from the 17th and 18th c.; www.brievenaltsbuit.nl), in order to determine how u could become a subject pronoun, and how the verbal agreement evolved with it. We argue that a very similar account to that proposed by Lara Bermejo (2016b) for the southern IR varieties is justified for the historical Dutch data as well.

2 Pilot study

2.1 Method

We searched the Letters as Loot corpus for forms of the verbs hebben ‘have’, zijn ‘be’ and zullen ‘shall’ and a form of U, UE, or UL, both in straight and inverted word order. In order to obtain clear results regarding the verbal agreement, we removed all the plurals (e.g., ue. zijn ‘you(r honourables) are’) because they are syncretic in all three persons. We further restricted the search to the verbs hebben, zijn and zullen because they do not have syncletic 2SG/3SG forms, as other verbs do), as in [(5)]. Morphologically heeft/zal is syncretic with 3SG, while hebt/zult is unambiguously 2SG, as shown in [6].

(5) jij loopt vs. hij loopt

4 We only focus on the verbal agreement, but note that there are indications that the possessive agreement may be worth looking at at a later stage, cf. example [5].
As we obtained too few data for all separate regions in corpus, we focused on the three regions with most data, Zeeland, Noord- and Zuid-Holland, and recoded the rest as “other”.

<table>
<thead>
<tr>
<th></th>
<th>17th c.</th>
<th>18th c.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2SG</td>
<td>3SG</td>
<td>total</td>
</tr>
<tr>
<td>Zeeland</td>
<td>7</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Zuid-Holland</td>
<td>4</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Noord-Holland</td>
<td>5</td>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>other</td>
<td>1</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>

### 2.2 Factors influencing the variation in verbal agreement

We analysed the data using multiple logistic regression in Rbrul ([Johnson 2009](#)), and found that region is the strongest predictor of verbal agreement with UE / UL / U, besides inversion (p = 0.0115). The model shows that Zeeland strongly favours 2SG agreement, while Zuid and Noord Holland prefer 3SG agreement. On the map, this results in a similar diffusion pattern to the one that [Lara Bermejo (2015, 2016b)](#) described for the agreement with ustedes in Western Andalusian dialects. It appears that the use of u with 3SG agreement was innovated in Holland and then diffused to Zeeland, where 2SG was initially more frequent, but is gradually replaced over time. Interestingly, this seems to suggest that in fact both theories regarding the origin of the nominative use of u may at least partially be right: on the one hand, u became used as a subject due to the (cross-linguistically common) shift from accusative to nominative, but initially kept the 2SG agreement. On the other hand, the 3SG agreement may still have arisen under the influence of the epistolary forms.

---

5Factor groups in the full model: region, period, gender, verb, inversion.
There is a second factor in the regression model that is significant at the $p<0.05$-level, viz. inversion. 3SG agreement is significantly more frequent in subject-initial V2 (“straight” contexts), as evidenced by the factor weight closer to 1 (0.723), while 3SG agreement much more likely in inversion contexts, as in (7) (factor weight closer to 0, viz. 0.277). We will return to this in Section 2.3 below.

(7) *dat sulke gedagten sijn nog nooijt of sullen nooyt in myn opkomen daar voor heeft Ue myn te veel goeds gedaan*
‘that such thoughts have never and never shall cross my mind, you have(3SG) done too much good for me, for such a thing to happen.’
(J.D. Piest to J.D. Praetorius, 1781/02/01)

The other factors (period, gender and verb) could not be shown to be significant.\(^6\) However, adjacency between the verb and the form of address may possibly play a role, though we have not been able to test this yet, due to restrictions of the search interface. Early examples of nominative *U* (i.e., not *UL/UE*) may indicate an influence, with non-adjacency correlating with 3SG agreement. Observe that in (9), there is even a 3SG possessive pronoun (*zijn*), besides the 3SG verb agreement.

\(^6\)We observed that *UL* is the form that most frequently occurs with 2SG verbal agreement, and that it is mostly used by women, but disappears after 1700. The addressee of the women’s letters is typically the husband or another family member, so the fact that 2SG sticks here for longer may indicate that despite the introduction of a new pronoun, the informal 2SG agreement remains used in informal contexts. *UE* is by far the most dominant form, and is most frequently used with 3SG verbal agreement.
2.3 Sketch of an analysis

In generative approaches, subject-verb agreement is analysed as a local dependency between a functional head, normally T (assigning nominative case to the subject), and the subject (checking T’s \( \phi \)-features). Typically, this dependency involves the subject moving to (or through) SpecTP. Assuming a symmetric view of V2, there is a second position for subjects in Dutch, viz. SpecCP. This is where the subject is found in subject-initial V2-clauses, while it remains in SpecTP in inversion contexts. In such contexts, SpecCP may be filled by a topic (amongst others).

In order to account for the variation in IR dialects, Lara Bermejo (2016b) argues that the replacement of vosotros by ustedes proceeds in three steps: (i) 3PL ustedes is merged as a topic (in SpecCP), doubled by the regular 2PL subject in SpecTP; (ii) 3PL ustedes in SpecCP is doubled by a covert 2PL clitic in SpecTP, triggering 2PL agreement on the verb; and finally, (iii) ustedes is reanalysed as a subject, and triggers 3PL agreement in SpecTP.

7 We assume a very simplified model here, with TP and CP serving as abbreviations for a possibly more fine-grained structure as would be assumed under a cartographic analysis.

8 Lara Bermejo uses Uriagerea’s (1995) “big DP” hypothesis for this, which Rubio Alcalá (2014) proposes to extend to account for clitic doubling, with the head of DP being filled by the clitic, and the complement by the topic.
We argue that this proposal in terms of doubling of a topic by a covert pronoun can be transferred to account for the variation in verbal agreement with subject $u$ in historical Dutch, as well. Our quantitative analysis suggested that 3SG is gaining on 2SG agreement, and that 2SG agreement is preferred in straight V2-clauses. Under our adaptation of Lara Bermejo’s proposal, 3SG $u$ is a full pronoun (in the sense of Cardinaletti & Starke) used as a topic in SpecCP in straight V2-clauses, and the 2SG agreement with the verb is mediated by a null pronoun in SpecTP.

\[(13)\quad [\text{CP} u [3SG] \quad [\text{C} [\text{zult}, [2SG]] \quad [\text{TP} \emptyset [2SG] \quad [\text{VP} \text{een wijf hebben}] \quad [\text{T}[2SG] t_i]]]\]

The assumption of a null doubling pronoun in this position is justified by the ample occurrence of overt pronouns in exactly this position, doubling a pronoun in SpecCP in (Southern) Dutch dialects, like *gie* in (14).

\[(14)\quad \text{Ge kent gie da.} \quad \text{(Haegeman & Van de Velde 2008: 163)}
\quad \text{‘You(2SG/Pl) know you(2SG) that.’}\]

In inversion contexts, $u$ appears in SpecTP, and triggers 3SG agreement on the verb.

\[(15)\quad [\text{CP daarvoor} \quad [\text{C} [\text{heeft}, [3SG]] \quad [\text{TP} \text{Ue} [3SG] \quad [\text{VP} \text{myn te veel goeds gedaan}] \quad [\text{T}[3SG] t_i]]]\]

Over time, the null pronoun in SpecTP was lost, and $u$ was reanalysed as a subject, with a trace in SpecTP in straight V2. 3SG agreement spread to some extent, helped along by the syncretism between 2SG and 3SG agreement found in most verbs. It is possible that the fact that epistolary forms of address when spelled out should trigger 3SG agreement, added to this.

\[(16)\quad [\text{CP} u [3SG] \quad [\text{C} [\text{zal}, [3SG]] \quad [\text{TP} t_i [3SG] \quad [\text{VP} \text{een wijf hebben}] \quad [\text{T}[3SG] t_i]]]\]

This proposal leads us to expect that embedded clauses, too, should show significantly more 3SG agreement in sentences with subject $u$ in 17th and 18th c. Dutch. This hypothesis is corroborated by the diffusion of agreement patterns with *ustedes* in the southern IR varieties, where third person agreement is found in embedded before main clauses (Lara Bermejo 2016a: 266-7). We leave this for fu-

\[9\text{This example is from the West Flemish dialect of Lapscheure, which has been extensively described and analysed by Liliane. Particularly the doubling of subject pronouns is one of her long-standing research interests.}\]
ture research. The fact that the variation between 2SG and 3SG agreement has stabilised in the present-day language seems to reflect a certain degree of lexicalisation: 2SG with *zijn*, (preferred) 3SG with *hebben*, for instance.

**Dedication**

With this contribution, we would like to thank Liliane for everything she has done for both of us, and keeps doing, for her generosity, always open door, and open ears throughout the years, for being a wonderful colleague, a mentor, a second mother, and a friend. Besides, she is one of the reasons there is an atmosphere within our research group DiAling that encourages fruitful collaborations such as this one, across language and framework boundaries.

**References**


Vor der Hake, J.A. 1911. Is de beleefdheidsvorm *U*n verbastering van *UEd.? *De Nieuwe Taalgids* 5. 16–24.


Lara Bermejo, V. 2016b. When agreement is for covert but not for overt: The case of ustedes plus second person plural inflections in Peninsular Spanish. *Isogloss* 2. 95–111.


The Haegeman test: 
A’-movement in the middle field

Hans Broekhuis

1 The test

As Liliane’s linguistic work is wide-ranging and of a uniformly high quality, I imagine that most linguists acquainted with it (and it is hard to conjure up any serious syntactician who has not read at least some of her publications) would find picking out a specific page from her work as their favorite an impossible task. However, I have one such page: page 179 of The Syntax of Negation (1995). This page shows that Dutch is just like West Flemish in that it has obligatory movement of negative phrases expressing clausal negation into the specifier of a NegP located in the middle field of the clause. Establishing movement within the middle field of the clause is often extremely difficult, as it can be and often is applied in a string-vacuous fashion; the two structures in (1), for instance, differ in the structural position of the noun phrase niemand ‘nobody’ but will nevertheless give rise to the same linear order because there is no phonetically realized material in between the two positions.

(1) a. dat Jan [NegP [Neg ø] [VP niemand ziet]].
   b. dat Jan [NegP niemand [Neg ø] [VP t i zien]].

   that Jan nobody sees
   ‘that Jan doesn’t see anyone.’

Liliane demonstrates in a very simple and elegant way that movement of the negative phrase is obligatory by means of what I have come to think of as The Haegeman test: when we embed a negative phrase in the PP-complement of a predicative AP, leftward extraction of the PP op niemand can be shown to be obligatory because the PP must precede the adjective. The fact that the PP must also precede the modifier erg ‘very’, which can be taken to be located in a high
position within the AP, supports the claim that the PP must in fact be extracted from the AP (2b); cf. *Jan is [AP erg op niemand dol].

(2) a. dat Jan [NegP [Neg ø] [VP [AP erg dol [PP op niemand]]] is ]]
    b. dat Jan [NegP [PP op niemand], [Neg ø] [VP [AP erg dol t ]] is ]]

    ‘that Jan isn’t very fond of anybody.’

The Haegeman test is crucial for determining the internal structure of the middle field of the clause, as well as the types of A′-movement available in this domain of the clause in the Germanic OV-languages. In other words, the test may help us determine the values available for X in structure (3) besides “Neg”.

(3) Clause-initial position                     Postverbal field
               ↓                          ↓
                [CP ..... C [TP ..... T [XP ..... X [VP ..... V ..... ]]]]]
                              Verb second & complementizer
                        Middle field                     Clause-final
                          position

The Haegeman test has proved of great value in the Syntax of Dutch: the discussion in Broekhuis & Corver (2016: §13.3) on A′-scrambling in the middle field of the clause could probably not have been written without it. The examples in (4) first show that Dutch has focus/topic movement within the middle field of the clause. Note that the use of italics indicates the so-called A-accent assigned to contrastive foci and the use of italics with underlining indicates the B-accent assigned to contrastive topics; see Neeleman & Van de Koot (2008), Broekhuis & Corver (2016: §13.3.2), and references cited there for detailed discussions of these notions.

(4) a. dat Marie [FocP [op PETER], Foc [VP [AP erg dol t ]]] is]
    that Marie of Peter very fond is
    ‘that Marie is very fond of Peter.’
    b. Ik weet niet wat Marie van Jan vindt, maar ik weet wel dat
    I know not what Marie of Jan considers, but I know AFF that
    ze [TopP [op Peter], Top [VP [AP erg dol t ]] is]].
    she of Peter very fond is
    ‘I don’t know how Marie feels about Jan but I do know she’s very fond

61
It is sometimes argued that focus movement differs from Neg-movement in that it is not obligatory because leaving the PP in its AP-internal position is possible if Peter is contrastively stressed: *dat Marie erg dol op PĆę Ćę is*; see Neeleman & Van de Koot (2008) and references cited there. There is, however, also evidence that the specifier position of FocP must be filled, based on the distribution of focus particles such as zelfs ‘even’. But first consider the examples in (5). (6b) shows that such particles may be left-adjacent to the focused phrase and (5b) shows that the particle and PP must be analyzed as a constituent because otherwise this example would violate the verb-second requirement on main clauses.

(5)  
(6a) *dat Marie [FocP zelfs op PĆę Ćę], Foc [VP [AP erg dol t.] is]].
that Marie even of Peter very fond is ‘that Marie is even very fond of Peter.’

(6b) dat Marie [FocP zelfs Foc [VP [AP erg dol [op PĆę Ćę],] is]].
that Marie even of Peter is Marie very fond ‘Marie is even very fond of Peter.’

The crucial fact is that while the focused phrase may occur within its original AP-internal position, the focus particle must be located in the specifier of FocP: see Broekhuis & Corver (2016: §13.3.2 sub IC2)) for a discussion of the question as to whether the focus particle is base-generated as part of the PP and moved into SpecFocP by focus movement, or whether it is inserted in SpecFocP directly by the merge operation; see Barbiers (2014) for independent arguments showing that focus particles of the zelfs type can occupy SpecFocP.

(6)  
(6a) *dat Marie [FocP _ Foc [VP [AP erg dol [zelfs op PĆę Ćę],] is]].
that Marie very fond even of Peter is ‘that Marie is even very fond of Peter.’

(6b) dat Marie [FocP zelfs Foc [VP [AP erg dol [op PĆę Ćę],] is]].
that Marie even of Peter is Marie very fond ‘Marie is even very fond of Peter.’

The acceptability contrast in (6) suggests that some form of focus movement is obligatory or, at least, that the specifier of FocP must be filled by some focus element. This of course raises the question why examples such as *dat Marie erg dol op PĆę Ćę is* are acceptable as well. The answer may be that such examples do not have an active FocP and thus should receive a different interpretation than examples with focus movement; I hope to return to this question in future work.
2 Intermezzo on the cartographic approach

Before continuing with the discussion of the Haegeman test, let me allow myself a brief cross-linguistic digression. Example (7a) shows that although English does not have focus movement of the kind found in Dutch, it does have a “low” FocP, the specifier of which can be filled by a focus particle: (7a) is unacceptable with a neutral intonation pattern (that is, without backgrounding or afterthought intonation). This is expected given that (7b) shows that the same also holds for Neg-movement; see Broekhuis & Klooster (2010) for relevant discussion. The examples in (6) and (7) therefore suggest that the availability of a “low” FocP is not a language-specific property of the continental Germanic OV-languages but a more general one.

(7) a. that Mary is even very fond of PETER.
     a’. that Mary is very fond even of PETER.
     b. that Marie is not very fond of anybody.

It is also worth pointing out that Broekhuis & Corver (2016: §13.3.2, sub IA1) provide evidence that the Dutch FocP is located in between the modal adverbs and NegP and that TopP is located higher than (that is: precedes) the modal adverbs (which goes against Neeleman & Van de Koot (2008), who deny the existence of such designated focus/topic positions). This would indicate that the Dutch middle field is structured in a similar way as the Hungarian preverbal field (cf. É. Kiss 2002). The fact that these genetically unrelated languages have the same linear order contrastive topic > contrastive focus > negation of course supports the cartographic approach to syntax, which has been one of Liliane’s main research interests over the last two decades.

3 The test as a linguistic tool

The Haegeman test is not only a crucial tool for establishing A’-movement in the middle field of the clause but can also offer invaluable help in evaluating analyses that avail themselves of such movements. A good example is the analysis of so-called correlative coordinators such asowel ... als ... ‘both ... and ...’, and of ... of ... ‘either ... or...’. The traditional analysis of such sequences is that we are dealing here with complex coordinators. Larson (1985), for instance, has proposed that either ... or ... originates as a single lexical head (Co) and that either is moved into some position preceding the first coordinand later in the
derivation, as illustrated in (8a). A useful piece of evidence in favor of Larson’s account is that we can now also derive examples such as (8b) by assuming that either may also target positions external to the coordinate structure.

(8)  
   a. that John ate [either, $[\text{CoP } \text{rice } [\text{Co} \text{t}_i \text{ or } \text{beans}]]$].
   b. that John [either, $[\text{VP } \text{ate } [\text{CoP } \text{rice } [\text{Co} \text{t}_i \text{ or } \text{beans}]]$]]).

It should be noted, however, that this argument does not immediately carry over to OV-languages such as Dutch because the verb follows the coordinate structure in examples such as (9), so that the presumed head movement in (9b) would apply string vacuously; the two structures therefore give rise to the same linear order (although Broekhuis & Corver, in prep., discuss a number of more complex Dutch examples illustrating the same thing). Note in passing that I assume that all Dutch correlative coordinate structures have the same underlying structure, an assumption that may not be true for English both … and … for reasons (related to the fact that it differs from Dutch zowel … als … in that it triggers plural subject agreement) that I cannot discuss here.

(9)  
   a. dat Jan [zowel, $[\text{CoP rijst } [\text{Co} \text{t}_i \text{ als } \text{bonen}]]$] at.
       that Jan both rice and beans ate
   b. dat Jan [zowel, $[\text{VP rijst } [\text{CoP } \text{rijst } [\text{Co} \text{t}_i \text{ als } \text{bonen}]] ]$] at.
       that Jan both rice and beans ate

Larson’s complex-head analysis of correlatives is fraught with problems for various reasons: one important problem (not mentioned in the literature as far as I know) is that the structures in (8) and (9) violate the lexical integrity hypothesis, which prohibits movement of a subpart of a lexical item: see Schwarz (1999) for a discussion of various other problems. An alternative approach to correlative coordinators is provided in Hendriks (2001, 2004) and Johannessen (2005), who argue that the initial part of a correlative coordinate structure is a focus particle. One argument in favor of this proposal is that the initial part of the correlative coordinate structure must have an emphatically accented phrase in its domain.

(10)  
   a. Peter heeft zelfs JAN ontmoet.
       Peter has even Jan met
   b. Peter heeft zowel JAN als ELS ontmoet.
       Peter has both Jan and Els met

Another argument in favor of this proposal is that it immediately accounts for
Neijt’s (1979) generalization that the coordinands in a correlative coordinate structure are normally major phrases (clausal constituents and certain verbal projections).

(11)  

a.  [Zelfs de mannen] waren aanwezig.  
    even the men were present  
    a’. *[De zelfs mannen] waren aanwezig.  
        the even men were present  

b.  [Zowel [de mannen als de vrouwen]] waren aanwezig.  
    both the men and the women were present  
    b’. *[De zowel [mannen als vrouwen]] waren aanwezig.  
        the both men and women were present

If correlative coordinate structures are indeed contrastively focused constituents, our earlier conclusion that the specifier position of FocP cannot remain empty predicts that either the full correlative coordinate structure or its initial element must be placed in SpecFocP. The Haegeman test shows that this prediction is indeed correct.

(12)  

a.  *Jan is [boos [zowel op Jan als op Marie]] geweest.  
    Jan is angry both at Jan and at Marie been  
    b.  Jan is [zowel op Jan als op Marie], [boos t.i.] geweest.  
    Jan is both at Jan and at Marie angry been  
        ‘Jan has been angry both at Jan and at Marie.’  
    c.  Jan is zowel [boos [op Jan als op Marie]] geweest.  
    Jan is both angry at Jan and at Marie been  
        ‘Jan has been angry both at Jan and at Marie.’

That the initial elements of correlative coordinate structures are focus particles can also be motivated by examples such as (13b), taken from Hoeksema (1989), in which a clausal correlative coordinate structure follows a clause-final verb (cluster); the acceptability contrast between the two competing word orders in (13b) is similar to that found in the competing word orders in the run-of-the-mill focus construction in (13a).

(13)  

a.  Jan heeft <alleen> gezegd <*alleen> dat Marie komt.  
    Jan has only said that Marie comes  
    ‘Jan has only said that Marie is coming.’
b. Jan heeft both gezegd dat Marie komt als dat Els komt.

‘Jan has both said that Marie is coming and that Els is coming.’

The examples in (14) also bear out that the same can be observed in the case of prepositional correlative coordinate structures (provided we do not use a back-grounding or afterthought intonation).

(14)  a. Jan heeft only gewacht op vader.

‘Jan has only waited for father.’

b. Jan heeft both gewacht op vader als op moeder.

‘Jan waited both for father and for mother.’

As a bonus, observe that example (15) shows that it is also possible to split the presumed correlative coordinate structure, which suggests that we are not dealing with coordination of clausal constituents at all; see {Broekhuis & Corver} for more discussion.

(15) Jan heeft zowel op vader gewacht als op moeder.

‘Jan waited both for father and for mother.’

The examples in (13) and (14) show that the Haegeman test is in fact applicable to all structures in which some element has to cross over another element with a fixed position in the clausal structure in order to reach some designated landing site. I am convinced that its scope will be expanded in future linguistic work and will ultimately become one of the standard devices in the linguistic tool kit.

Acknowledgement: I would like to thank Frits Beukema for copy-editing this contribution and Liliane for the many pleasant, interesting discussions we have had over the last two decades.
References


Early on in the generative literature some phenomena that were considered to be limited to root clauses have been identified (cf. Emonds 1969 and Hooper & Thompson 1973). As Aelbrecht et al. (2012) note in their assessment of this literature, phenomena that have been traditionally analyzed as occurring only in main clauses come in at least two varieties: those that are root phenomena in a strict sense and those for which the characterization as root phenomena might be questioned. For example, a phenomenon that was initially treated as restricted to the root is argument fronting in English, based on contrasts like the one in (1)–(2).

(1) This book you should read.
(2) *It is impossible that this book he has read.

However, as already noted in the early literature cited above, argument fronting is possible also in clausal complement of verbs of saying. Although these cases might not be strong counterexamples since clausal complement of verbs of saying might be sufficiently “root-like”, subsequent work by Liliane Haegeman (cf. Haegeman 2012 for a systematic presentation) has shown that argument fronting is possible also in certain types of adverbial clauses, those which Haegeman calls peripheral adverbials (cf. Haegeman 2003). An example is (3).

---

*I am very happy to contribute this squib to Liliane's Webschrift. Liliane, as a teacher and as a person, was crucial in getting me into linguistics. Before enrolling in her class on negation in the Fall 1992 at the University of Geneva, linguistics, and syntax in particular, was just a small curiosity in an agenda dominated by other interests. But everything changed during that year in Geneva and Liliane, a wonderful teacher and an extraordinary researcher, played a big role in that (whether this was for better or for worse is a different matter, but Liliane cannot be considered responsible for that!).*
While YOUR book they are using in two courses, mine they haven’t even order for the library. (Haegeman 2003: 332)

For “root phenomena” like argument fronting, Haegeman developed an account based on intervention which does not make a crucial use of the distinction between root and non-root clauses. This seems a welcome fact given that these phenomena do not occur only at the root, despite the initial characterization in this direction.

However, there are genuine root phenomena. A well-known case is V2 in Dutch and German. Less well-known cases include hanging topics in Italian (cf. Cinque 1977) and right dislocation in “strict OV languages” like Japanese and Turkish (cf. Cecchetto & Donati 2015 for an account that unifies these three phenomena and tries to explain why they are restricted to main clauses). My limited goal in this squib is to discuss a new case of root phenomenon in Italian that, as far as I know, has not been previously systematically studied. This can be illustrated by the question-answer pair in (4):

(4) E il ragazzo? Partito
And the boy? Left
‘What about the boy?’ ‘He left.’

The answer in (4) is interpreted as a declarative tensed clause although it is very reduced, the only overt element being the past participle partito. Neither the auxiliary nor the only argument of the verb are overtly expressed. For concreteness, I will use the label “internal argument drop sentences” to refer to the construction exemplified in (4), but I want to stress that in addition to the internal argument the auxiliary is dropped as well. The choice of the label is motivated by the fact that the construction is attested only with internal arguments. This is confirmed by sentences (5) to (7): (5), which contains a passive, is fully acceptable, much like (4), which contains an unaccusative verb. However, (6), which contains an unergative verb, and (7), which contains a transitive verb, are sharply ungrammatical.

(5) E il dessert? Mangiato (da Leo)
And the dessert? Eaten (by Leo)
‘What about the dessert?’ ‘Leo ate it.’

(6) E il ragazzo? *Pianto
And the boy? Cried
In Italian, the past participle agrees in gender and number with the internal argument of passive and unaccusative verbs. This extends to internal argument drop sentences, where agreement is with contextually given internal argument. While in (4) and (5) the make-up of the past participle is compatible with its default value (singular, masculine), (8) and (9) show that the past participle is inflected in gender and number.

(8) E le torte? Mangiate (da Leo)
And the cake-PLUR-FEM? Eaten-PLUR-FEM (by Leo)
‘What about the cakes’ ‘Leo ate them.’

(9) E le ragazze? Partite
And the girls-PLUR-FEM? Left-PLUR-FEM
‘What about the girls?’ ‘They left.’

Internal argument drop sentences are not restricted to question-answer pairs, as long as the dropped argument is contextually salient. For example, imagine a context in which I enter my office and I notice that the desk next to mine has been fully emptied. A colleague might react to my puzzled look by uttering (10).

(10) Licenziata dal capo.
Fired-SING-FEM by the boss
‘She has been fired by the boss’

Internal argument drop is not allowed in a declarative with a fully-fledged verb, as shown in (11). For internal argument drop to be possible, the verb must be a past participle.

(11) *E quel film? vedrò (domani)
And that movie? (I) will-watch (tomorrow)

It should be clear that the internal argument drop construction is not a simple case of ellipsis to be analyzed only at the discourse level but obeys very specific syntactic constraints and thus require a syntactic analysis. In fact, the constraints on the internal argument drop construction are reminiscent of those governing another reduced construction in Italian, namely the one called absolute small clause by Belletti (1999). An absolute small clause is an adjunct clause composed by a past participle which agrees with an internal argument DP, which can remain
unexpressed:

(12) Partite (le ragazze), restammo soli Left-PLUR-FEM (the girls-PLUR-FEM) remain-PAST-1-PLUR alone
‘After the girls left, we remained alone.’

(13) Mangiata la torta, facemmo una passeggiata Eaten-FEM-SING the cake-FEM-SING make-PAST-1-PLUR a walk
‘Having eaten the cake, we took a walk.’

An obvious analogy between absolute small clauses and the internal argument drop construction is that they are both restricted to past participles that combine with an internal argument (also absolute small clauses are restricted to passive and unaccusative verbs). But the analogies do not stop here. As Belletti (1990) discusses, negation is not allowed in absolute small clauses (cf. (14)). (15) shows that the same holds for the internal argument drop construction.

(14) *Non mangiata la torta, facemmo una
NEG eaten-FEM-SING the cake-FEM-SING make-PAST-1-PLUR a
passeggiata
walk

(15) E le ragazze? *Non partite
And the girls-PLUR-FEM? NEG Left-PLUR-FEM

Belletti (1990) also notices that, although ne-extraction is possible from the internal argument of a transitive verb (cf. (16)), ne-extraction is not possible in absolute small clauses (cf. (17)). The same holds for internal argument drop sentences (cf. (18)).

There is just one context that licenses a negated past participle in isolation and this is enumeration in a list. For example, in the old days of the Italian weather broadcasting a list of cities would be read and the minimum/maximum temperature registered in each city the previous day would be given. In those radio broadcastings, the frozen expression non pervenuta (‘not arrived’) was used, as shown in (i):

(i) Bari 20-28
    Bolzano non pervenuta
    Milano 18-28
    Roma 20-30
    Etc.
Despite these analogies, there is a striking difference between absolute small clauses and internal argument drop sentences: absolute small clauses, being adverbial clauses, need to be embedded, while the internal argument drop construction can never be embedded. We can check this by trying to embed it under a verb of saying. The result is ungrammatical (cf. (19)).

The same ungrammaticality is observed when the internal argument drop construction is embedded in a peripheral adverbial clause (i.e., the structures that Haegeman has shown to allow argument fronting). This is shown in (20) and (21). (21) is a telling case, as it contains a premise conditional, which is prototypical example of peripheral adverbials.

We can conclude that internal argument drop is rigidly a root phenomenon. An even more striking observation is that internal argument drop sentences, despite being very reduced (in fact, they typically contain only a past participle) have a full force specification. They can be declaratives as in the example discussed up to now, but they can also be interrogatives (cf. (22)) or exclamatives (cf. (23)).

We can conclude that internal argument drop is rigidly a root phenomenon. An even more striking observation is that internal argument drop sentences, despite being very reduced (in fact, they typically contain only a past participle) have a full force specification. They can be declaratives as in the example discussed up to now, but they can also be interrogatives (cf. (22)) or exclamatives (cf. (23)).
‘Has he been fired?’

Context: Leo enters his office, notices that the desk next to his has been fully emptied and utters (22) to elicit information from a colleague.

(23) Bruciata!
Burned-SING-FEM

Context: Leo arrives next to the Opera House, notices that it has been destroyed by a fire and utters (23) to express his surprise/ disappointment.

Internal argument drop sentences raise several interesting questions. One is the nature of the dropped argument. A possibility is that it is pro, somewhat licensed by the agreement morpheme on the past participle (cf. Rizzi 1986). A more complex issue is the following. On the one hand, internal argument drop sentences share many properties with small clauses: they cannot contain an external argument, they are not compatible with a full verbal morphology (either in the form of an auxiliary or in the form of a finite verb) and they cannot be negated. In structural terms, this seems to suggest that they are reduced structures, possibly smaller than vP (assuming that Spec,vP is where the external theta role assigned, cf. Chomsky 1995 and Kratzer 1996). On the other hand, internal argument drop sentences have a full force specification and the projection ForceP is taken to be the highest projection in the CP area (cf. Rizzi 1997 and much following work).

Confronted with this puzzle there are at least two ways to go. Either we assume that ForceP can be projected in the lower segment of the structure, say in the VP periphery. Or we can assume that the intermediate projections between VP and ForceP can be deactivated. It is not even clear that these two proposals are not notational variants of each other, since they both boil down to saying that an extended middle portion of the structure can be absent under certain conditions.

But the real issue is explaining why this middle portion can be missing in internal argument drop sentences and (apparently) only in them. Although I am not in the position to answer this question, I hope to have shown that internal argument drop sentences are challenging enough to deserve a serious investigation.
References


A Note on Relative Clauses with Split Antecedents

Guglielmo Cinque

Since Perlmutter & Ross (1970), relatives with split antecedents have represented an analytical problem for any theory of relative clauses. Two cases should in fact be distinguished; one in which the split antecedents occur in two (or more) co-ordinated sentences, characteristically bearing the same grammatical function (subject, object, etc.) (see section 1) and one in which they bear different grammatical functions in one and the same sentence (see section 2).

1 Split antecedents in coordinate sentences

The examples in (1) illustrate the case of split antecedents in two coordinated sentences:

(1) To Liliane with sympathy and admiration.

This construction must be kept distinct from the construction with coordinated antecedents which has come to be known as the hydra construction Link (1984). The reason is that languages with pre-nominal relatives have the latter but not the former.

2 Note that (1)b., e., f. and g. have split antecedents in the object rather than in the subject position of the coordinate sentences (pace Rochemont & Culicover 1990: 38f). More difficult seem cases where the split antecedents in the coordinated sentences have different grammatical functions. Baltin (2005: 255) gives an example like (i) as ungrammatical:

(i) *A man entered the room and I saw a woman who were similar

Also see Moltmann (1992). Yet, as noted in Smits (1989), for some Italians the split antecedents of a nonrestrictive relative need not have the same grammatical function. In fact I tend to accept a sentence like (ii):

(ii) Alla fermata è arrivato un uomo ed io ho visto anche un ragazzo, i quali si assomigliavano molto

‘At the bus stop a man arrived and I saw a boy too, who looked very much alike.’
(1)  a. A man entered the room and a woman went out who were quite similar. (Perlmutter & Ross 1970: 350)
   b. Every villager envies a relative of his and every townsman admires a friend of his who hate each other (Hintikka 1974: 172)
   c. The girl left and the boy arrived who met in Chicago (Chomsky 1975: fn.47)
   d. It is obvious that a man came in and a woman went out who were similar (Andrews 1975: 119)
   e. John saw a man and Mary saw a woman [who were wanted by the police] (Alexiadou et al. 2000: 14)
   f. Kim likes muffins, but Sandy prefers scones, which they eat with jam (Arnold 2004: 274)
   g. John noticed a man and Mary spotted a woman [who the police were looking for –] (Radford 2017: §5.2)

Our tentative interpretation of such cases is that they are possibly to be assimilated to those RCs that (marginally) receive a restrictive interpretation even though they enter a discourse grammar nonrestrictive structure like the restrictives with heavy pied piping in English (as in examples like every candidate the father of whom Bill voted for — Jacobson 1998: 81) or the restrictives employing the art. + qual- paradigm in Italian (i soci i quali non abbiano ancora versato la quota annuale.. ‘The members who have not paid the annual fees.. ’— Cinque 1982: 264); constructions not derived by raising, in which the wh-pronoun is interpreted similarly to a pronoun or demonstrative. For additional discussion of the discourse grammar non-restrictive construction, also arguably used for the special restrictive construction examined here, see Cinque (2008, to appear).

This interpretation may be supported by the following three facts. First, replacement in English of a wh-relative pronoun with that (which is otherwise un-

Chaves (2012: §3.4.3) notes that conjunction, but not disjunction, gives acceptable sentences:

(iii) *A man entered or a woman left who were quite similar.

Chomsky (1975: 98) actually says “To me these examples seem at best quite marginal, and I would question whether anything can be based on them.” (fn.47). In fact not all languages appear to allow for them. See Cardoso (2010: 191f) on European Portuguese.

Recalling the analysis of split antecedents of nonrestrictive relatives in Demirdache (1991: 116). Also see Yoshitaka Erlewine & Kotek (2013) and Webelhuth et al. (2013: 47), where such cases as (1a) are suggested to be similar to A man entered the room and a woman went out. They were quite similar.
exceptionable with embedded and extraposed restrictives and marginal to impossible with nonrestrictives) leads for at least some speakers to much less acceptable sentences. Megan Rae (p.c.) finds (1a) with that replacing who much worse (see (2a) with her judgment), and Arnold (2004: 30) marks the variant with that in (2b) as unacceptable:

(2)  
   a. ?*A man, entered the room and a woman went out that, were quite similar.  
   b. Kim likes muffins, but Sandy prefers scones, which, they eat with jam.

Second, an example like (3) in Italian is to my ear marginally possible even though it apparently violates the Right Roof Constraint:

(3)  
   ?[Che [qualcuno, ci abbia aiutato] e [un’altro, si sia aggiunto] è una fortuna] senza i quali this would not have been possible.  
   ‘That someone helped us and someone else joined in was a stroke of luck without whom this would not have been possible.’

Third, as noted in Del Gobbo (2010: 406f, 2015: §2.2, 2017: §2.2) and Lin & Tsai (2015: 105f) split-antecedent RCs parallel to (1) above appear not to be possible in Chinese, even in nonrestrictive RCs, which are of the integrated type. This may well be a general property of languages with pre-nominal RCs, which, as seen, do not dispose of non-integrated nonrestrictives (Jaklin Kornfilt, p.c., informs me that examples like (1) above are indeed impossible also in Turkish pre-nominal RCs, although examples of coordinated antecedents (hydrams) are perfectly grammatical).

For further discussion and different analyses of these split antecedent cases in terms of movement see de Vries (2002: 66ff and Chapter 7, §5.2.12), Zhang (2007, 2010), Cecchetto & Donati (2015: §3.3.5), Overfelt (2015: §6.2) and Fox & Johnson (2016). Despite these attempts, it does not seems unreasonable to conclude, with Alexiadou et al. (2000: 14), that “[w]hile it is feasible for an RC to be linked to multiple antecedents by a rule of construal, as in the standard

\(5\) Overfelt notes that examples like (i) below suggest, contrary to these expectations, that a negative polarity item can be licensed in the extraposed material even given split antecedence:

(i)  
   \([DP \text{ Every intern }]_1 \text{ left and } [DP \text{ every employee }]_1 \text{ quit } [CP \text{ who were in any of the basement offices}]_1 \).
approach, to claim that they are linked by a movement dependency is problematic. It seems rather far-fetched to suppose that the antecedents in [(1)] could have originated inside the relative clause (say, as a conjoined DP) to then be split and distributed across two clausal conjuncts after raising (a kind of ‘reverse’ Across-The-Board raising).”. Also see Andrews (1975) and McKinney-Bock (2013), McKinney-Bock & Vergnaud (2014), Radford (2017: §5.2) for non raising analyses, and Baltin (2005).

2 Split antecedents with different grammatical functions belonging to the same sentence

The cases of RCs with split antecedents belonging to the same sentence appear to be possible in Italian and English as nonrestrictive RCs. See (4a) and (4b), but not as restrictives. See the ungrammaticality of (5a)–(5e).

(4)  a. Se Carlo, non amava più Anna, i quali d’altra parte non si erano mai voluti veramente bene, una ragione c’era. ‘If C. was no longer in love with A. that at any rate never really loved each other, there was a motive.’
    b. Se Piero, non si trova più tanto bene con Ida, tra i quali d’altronde non c’è mai stata una vera amicizia... ‘If P. no longer likes to stay with I. between whom in any event there never was a real friendship, ...’

(5)  a. *The dog is chasing the cat which were fighting (Andrews 1975: 116)
    b. *A man met a woman yesterday who were similar (Guéron 1980: 648; credited to N. Chomsky)
    c. *The boy looked at the girl who both like sports. (de Vries 2002: 67)
    d. *A man visited a woman (yesterday) who were similar (Baltin 2005: 255)
    e. *Il ragazzo guardava la ragazza che entrambi amano gli sport (same as (5c))

Yet to judge from Chomsky (1975: fn47) referring to what would later be published as ex. (26) in (Perkins 1982: 284) similar sentences are apparently possible in Navajo (also see the discussion in Andrews 1975: 116ff):
(6) Łééchąą mósi yinoołchééł ahigánę́'ę́. (Navajo; Perkins 1982: 284)

*The dog_{i} is chasing the cat_{j}, which_{i,j} were fighting*.

In fact, they appear possible even in English, provided that the two antecedents are related by a symmetric predicate. See Poschmann et al. (2016), citing an example, (7) from Hoeksema (1986: 69):

(7) We always let those boys_{i} play with those girls_{j} [who_{i,j} know one another from elementary school].

Once again, such cases of split antecedents are impossible in languages with prenominal relative clauses. See Del Gobbo (2010), Del Gobbo (2015) on Chinese.

A raising analysis for this second type of split antecedent relatives would again seem to require special assumptions, while the same discourse grammar analysis of nonrestrictives mentioned above appears to be able to provide an analysis for this second type of split antecedent relatives.

References


Lin, J.W. & Tsai, W.T.D. 2015. Restricting non-restrictive relatives in Mandarin


A dummy and a diminutive in Dutch verbal lore

Norbert Corver

1 Introduction

In her study of subject omission in present-day written English, Liliane Haegeman draws our attention to the theoretical relevance of peripheral data for linguistic theorizing (see also Schmerling 1973, Thrasher 1977). Specifically, English (core grammar) does not allow subject omission in finite clauses, as shown in (1), but in specific written registers such as diaries, personal letters, e-mails or notes, English does allow for a subject to be non-overt. This is exemplified in (2), examples drawn from Haegeman (1990, 2007).

(1) a. *Have bought a book.
   b. I have bought a book.

(2) a. ∅Have done 110 pages. (Diary of Virginia Woolf, p. 33; 11 November)
   b. ∅Dreamt that I picked up a New Yorker. (Plath 304)

In this article I discuss a peripheral phenomenon attested in Dutch nursery rhymes and children’s songs (so-called verbal lore), namely the pattern van je XP, where XP can be a linguistic expression that designates a sound-symbolic, sequential (e.g. counting) or repetitive activity. The Dutch children’s song in (3) displays four instances of this phenomenon, each of which is marked in boldface. The left part in (3) represents the Dutch nursery rhyme, while the right part gives the (literal) English translation for each line of the Dutch verse.

---

The nursery rhymes and children’s songs can be found in the following data sources: van Vloten & Brandts-Buys (1894), Abramsz (1911) and the following website: http://www.overtuin.net.
The article is organized as follows: Section 2 examines the grammatical nature of *van* in expressions such as *van je ras ras ras*, and section 3 does the same for the element *je*. It will be proposed that *van* is a dummy (i.e., meaningless) element that spells out the categorial node *no* and that *je* is a diminutive morpheme, which appears to be able to occur on its own; that is, it does not need a lexical category – a noun – to which it can attach. Section 3 discusses a few other structural environments in which this “independent” diminutive morpheme possibly is present. Section 4 concludes the article.

2 Van + XP in Dutch verbal lore

Before turning to the pattern *van je XP*, I discuss a different but arguably related pattern, viz. *van + XP*. Some illustrations are given in (4):
(4) Daar ging een meisje over het land. there walked a girl across the land
Zij had een korfje al in haar hand. she had a basket-DIM PRT in her
hand
Daar was in van gladderdegla ď. there was in of smooth-er-de-
smooth
Daar was in van strijkerdestrijk. there was in of pet-er-de-pet
Daar was in van kijkerdekijk. there was in of look-er-de-look
Daar was in van krolderdekrol. there was in of caterwaul-er-de-
caterwaul
(riddle rhyme)

Twee kinderen zouden naar school gaan. two children would to school go
Ze waren zo vrolijk en blij. they were so cheerful and happy
In de verte hoorden ze een orgeltje. in the distance heard they a barrel-
organ-DIM
Daar moesten ze eventjes bij. there had.to.stand they briefly PRT
En ‘t orgeltje speelde van holie ha hij. and the barrel-organ-DIM played of
holie ha hij
En de kinderen dansten er bij. and the children danced there with
(= while the organ played)

The linguistic expression van gladderdegla ď fulfills the role of (argumental) sub-
ject of a clause. The sentence Daar was in van gladderdegla ď can be paraphrased
as “something very smooth was in it, i.e. in the basket”. The other van+XP ex-
pressions fulfill the same role and have a similar meaning: van strijkerdestrijk
designates something strokeable, van kijkerdekijk something which is looking
around, and van krolderdekrol stands for something which can caterwaul. Ar-
guably, the expression van holie ha hij in the second nursery rhyme fulfills the
role of direct object. The sentence containing it can be paraphrased as: “And
the barrel-organ played this”, where this corresponds to “holie ha hij”.

The question obviously arises how to analyze van. What I would like to pro-
pose is that van ‘of’ is a minimal manifestation of the syntactic category ‘noun’. In
a way, it acts like a dummy noun, whose contents is provided by the expression
that follows van (e.g., gladderdegla ď). In what follows, I will briefly give some
background for this analysis of van, basing myself on earlier generative-linguistic
analyses of English of.

In Chomsky (1986), the element of is considered to be a manifestation of
As Chomsky (p. 194) notes, “Genitive Case is realized morphologically by affixation of some element to the NP: of in complement position, the possessive element POSS [= ‘s; NC] in subject position.” These two modes of morphological realization are depicted in (5).

(5)  a. the [destruction [the city]] → the [destruction\textsubscript{N} [of + [the city]]]
    b. [[the city] destruction] → [[the city] + ‘s] destruction]

In Emonds (1985) and Pesetsky (2013), it is argued that Case is not a primitive category but rather an affixal realization of a Part-of-Speech. In other words, Case is a part-of-speech-suffix, or in Emonds’s terms: an ‘alternative realization’ of a categorial head. According to this approach, genitival Case is an affixal realization of the category ‘noun’. More specifically, of and -s in (5) are nominal affixes “assigned” by the noun destruction to the satellite constituent the city. To phrase it differently, the nominal property (i.e., N) associated with destruction is realized alternatively (i.e. affixally) on the satellite phrase. Schematically, for a phrase like the destruction of the city:

(6)  the [destruction [the city]] → the [destruction\textsubscript{N} [N\textsubscript{aff} (= of) + [the city]]]

With Pesetsky (2013), I take there to be two ways in which Case can appear on a constituent: (i) syntactic case assignment, as, for example, in (6), and (ii) Case as a lexical property. As regards this last way of Case appearance, Pesetsky (2013: 8) makes the following statement:

“[..] every element that comes from the lexicon as a noun, determiner, verb or preposition could equally well be described as coming from the lexicon assigned to the corresponding case-categories. In other words, from the point of view of syntax, every noun can be described as ‘born genitive’, every verb as ‘born accusative’, every determiner as ‘born nominative’, and every preposition as ‘born oblique’.”

According to this statement, one should be able to find overt manifestations of genitival Case (i.e., affixaal N) on nouns. I will argue that this is exactly what we find with van in (3) and (4). That is, van is the manifestation (spell-out) of the “genitival property” with which N is born. In Pesetsky’s (p. 8-9) terms, van is a surface manifestation of “the noun’s primeval state – that is, the form in which it entered the syntactic derivation [...]”. But before elaborating on this, I would like to make one more theoretical step. In line with Marantz (1997), Borer (2005) and
others, I assume that lexical categories (nouns, verbs, et cetera) have the form f-Root, where f is a categorial head and the Root (henceforth √) is unspecified as to category. Thus, the English noun car and its Dutch counterpart auto have a composite structure: \([n_P ~ n^\circ \[\sqrt{P} \sqrt{\text{car/auto}}]\])

For the above-mentioned approach to Case, this means that genitival case is an affixal no. This categorial affix can surface on a satellite constituent of the noun through case assignment, or it can surface on no itself (the “no-as-born-genitive” way; i.e. primeval genitive). The latter strategy can be represented as follows: \([n_P ~ n (= \text{van}) [\sqrt{\text{Root}}]]\).

The question obviously arises whether there are any nominal expressions displaying the “no-as-born-genitive” strategy. Notice that we do not find van in the following nominal expressions: \([\text{DP} \.de \[n_P ~ (\neq \text{van}) [\sqrt{\text{auto}}]\]]\); (the (of) car; intended meaning: ‘the car’). Possibly, this relates to the fact that the nominal nature of the phrase is already clear from the presence of the definite article de. Furthermore, raising of the Root auto to no yields the amalgam \([\sqrt{\text{auto}+n^\circ}]\). The overtness of the Root may make the appearance of van superfluous and, for reasons of economy, impossible. That is, I take the surfacing of n^\circ as van to be a last resort strategy, just like English do-support (Chomsky 1957, Lasnik 2000), which is found, for example, in clausal environments featuring a silent verbal complement of T (John did ∅ too) or an overt pro-form that substitutes for the verbal complement (John did so too). In short, van, just like the auxiliary to do, is a dummy element that surfaces in order save a structural representation.

I take the element van in (7)-B’ to instantiate the “born-genitive” strategy. That is, van represents a nominal expression of the following type: \([n_P n (= \text{van}) [\sqrt{\text{Root}}]]\). More specifically, I take it to be an indefinite pronominal phrase which, just like its definite counterpart ‘t ‘it’ (i.e. \([\text{DP} \ ‘t [n_P ∅]\]) in (7)-B, has the proposition ‘Jan passes the exam’ in (7)-A as its antecedent.

(7)  
A: Haalt Jan het examen?  
    passes Jan the exam  
    ‘Will Jan pass the exam?’

B: Ik denk [DP ‘t], wel/niet t.  
    (‘t = proposition ‘Jan passes the exam’)  
    I think [] it for.sure /not  
    ‘I think so / I don’t think so.’

B’: Ik denk \([n_P \text{van}], \text{wel/niet} \quad t.\)  
    (\text{van} = ‘Jan passes the exam’)  
    I think of for.sure/not

\(^2\)Interestingly, van in (7)-B’ requires the presence of the polarity marker wel/niet. Thus, the string ik denk van is ill-formed. The pro-form ‘t, on the contrary, cán occur without wel/niet, as in: Ik denk ‘t (I think it, ‘I think so’).
‘I think so / I don’t think so.’

Let us now return to the linguistic expressions *van gladderdeglad* and *van holie ha hij* in (4). I propose that the element *van* is a surface manifestation of the categorial head no. I tentatively propose that its appearance relates to the special status of the Root. That is, the Root designates something that is not directly associated with the conceptual meaning of a noun (say, denoting an entity or individual). For example, *holie ha hij* has sound-symbolic meaning, *gladderdeglad* designates an attribute (‘being smooth’), and *kijkerdeijk* designates an event/activity (‘to look’). Possibly, the inner structure of these expressions—maybe, a coordinate structure consisting of coordinated roots; see Corver (2014, 2015) —blocks raising of the Root to n°. As a result of the special nature of the Root, n° surfaces as *van* in order to make the nominal nature of the entire expression recoverable. Thus, *van gladder de glad* and *van holie ha hij* have the structures in (8).³

\[
\begin{align*}
(8) & \\
& \text{a. } [np \text{n° (= } \text{van})] [\text{Conjp } \text{gladder} [\text{Conj'} [\text{Conj de} ] \text{glad}]] \\
& \text{b. } [np \text{n° (= } \text{van})] [\text{Conjp } \text{holie} [\text{Conj'} \text{Conj'} \text{ha]} [\text{Conjp } \text{Conj'} \text{ha hij}]]] \\
\end{align*}
\]

The next section discusses the pattern *van je ras ras ras*, which was introduced in (3). The question that needs to be answered is: what is the grammatical nature of *je*?

### 3 Van je XP

Before the question regarding the nature of *je* is addressed, it is useful to give some additional examples of children’s songs containing the pattern *van je XP*.

³I take *de* to be decomposable into *d* and *e*. The latter element is a minimal spell-out (viz., schwa) of the Conj-head (see Corver 2014, 2015). The presence of *d* presumably results from a phonological rule: insertion of /d/ in the phonological environment *t* _er_. Compare insertion of /d/ when the agentive suffix *-er* is attached to a verbal root: *verhuur* ‘to let’, *verhuur-d-er* ‘landlord’.

⁴The representation *Conj<sub>∅</sub>* stands for a phonologically empty coordinate conjunction.
In Holland staat een huis (2x)
In Holland staat een huis, ja, ja.

Van je singela singela hopsasa
In Holland staat een huis (2x)

Drie schuintamboers die kwamen uit het oosten (2x)
Van je rom bom, wat maal ik erom?
Zij kwamen uit het oosten, rombom

En van je hela hola,
houdt er de moed maar in! (3x)
En van je hela hola
houdt er de moed maar in! (2x)

En van je hotsie knotsie knetter,
van je jippie jippie jee.
Maar zijn paard was zeer vermoeid
die wou niet verder mee.
Maar hij moest de boeven vangen
dus nam hij een ander beest
en nu mag je zelf bedenken
wat voor beest dat is geweest

The first hypothesis that comes to mind with regard to je’s grammatical nature is
the following: je is a second person singular weak possessive pronoun (see Hae-
seryn et al 1997). Such a pronoun normally appears in possessive noun phrases
such as je adres in (10).

(10) Is dit je adres?
is this your address
‘Is this your address?’

The question obviously arises whether there is any support for such a possessive
pronominal analysis. If it is a second person pronoun, one might try to connect
its appearance to the availability of an addressee (the hearer/reader) in the dis-
course context. The sequence *En van je hela, hola, houdt er de moed maar in!* is interesting in this respect, since *van je* *XP* precedes an imperative clause. As has been shown in the literature on Dutch imperatives, there is a silent pronominal subject (say, *YOU*) present in the imperative clause; see e.g. [Bennis (2006)](Bennis2006). It is imaginable that *je* in *van je* *XP* is somehow connected to the second person pronominal subject of the clause. Schematically:

\[(11) \text{and of je hela hola keep there the courage PRT in} \]
\[
\text{and of je hela hola keep there the courage PRT in} \\
\text{`And hela hola, keep it up! / stay positive!'}
\]

Clearly, this argument from imperative clauses does not have much strength, since many other clauses preceded by *van je* *XP* do not contain a second person subject, but rather a first or third person subject.

\[(12) \]
\[
a. \text{Van } je_{2PSG} \text{ rom bom, wat maal } ik_{1PSG} \text{ erom?} \]
\[
\text{of } je \text{ rom bom why whine I about it} \\
\text{`Boom boom, why whine about it?'}
\]
\[
b. \text{Van } je_{2PSG} \text{ ras, ras, ras, rijdt de koning}_{3PSG} \text{ door de plas.} \\
\text{of } je \text{ ras ras ras goes the king through the puddle} \\
\text{`Go go go, there goes the king through the puddle!'}
\]

Notice also that the possessive pronominal analysis faces a number of problems: First of all, it is not clear at all what exactly the possessive relationship is between, for example, *je* and the sound symbolic sequence *hela hola*. Secondly, the weak possessive pronoun cannot be replaced by its strong counterpart *jouw*: *Van jouw ras ras ras*. If *je* is a second person possessive pronoun, it is unclear why the strong pronominal form is impossible.

Instead of claiming that *je* is a possessive pronoun, I tentatively propose that *je* is a diminutive morpheme. Thus, *je* corresponds to the element that we normally find attached to a nominal host, as in

\[(13) \]
\[
\text{Ik heb } vlecht-je-s \text{ in mijn haar.} \\
\text{I have braid-DIM-PL in my hair}
\]

At first sight, this diminutive analysis does not seem very plausible. For one thing, *je* cannot occur independently; that is, it needs a nominal host to which it can be attached:
(14)  a. *Ik heb *een *je in mijn haar.
    I have a DIM in my hair
    ‘I have something small in my hair.’

    b. *Ik heb *jes in mijn haar.
    I have DIM-PL in my hair
    ‘I have small things in my hair.’

Even though it is true that the diminutive morpheme normally needs a nominal host to which it can attach, there are structural environments in which the diminutive -je does not combine with a noun, at least not with an overt noun (cf. Corver (to appear)). Consider, for example, the italicized patterns in (15):

(15)  a. Jan reed [erg zacht-je-s]
    Jan drove very slow-DIM-s
    ‘Jan drove very slowly.’

    b. [Hoe zacht-je-s] reed Jan?
    how slow-DIM-s drove Jan
    ‘How slowly did John drive?’

In these examples, je is directly preceded by an adjective and followed by the bound morpheme -s. This -s must be present: *erg/hoe zachtje. The phrase erg/hoe zachtjes in (15) has an adverbial function; it designates the manner in which Jan’s driving took place. Although an analysis according to which erg/hoe zachtjes is an adverbially used adjective phrase is tempting, it faces the problem that the diminutive -je normally does not attach to adjectives. It typically combines with nouns. Sticking to the generalization that -je only attaches to nouns, I propose an alternative analysis for erg/hoe zacht-je-s, namely the one given in (16):

(16)  [FP [erg/hoe zacht] [F: F [Clasp -je [nP n° (= -s) [√P √WAY]]]]]

According to this analysis, erg/hoe zacht is an attributive adjective phrase contained within a nominal expression whose Root is silent. As indicated by WAY, I take this silent root (Kayne 2003) to designate manner. Following Wiltschko (2005), I analyze the diminutive morpheme as a classifier that conveys ‘small piece’ (see also De Belder 2011). In informal terms, -je + WAY designates ‘small

5 The silent noun TIME is also possible in this structural environment: even-tje-s (just-DIM-s, ‘just, a little while’).
manner”, which possibly is at the basis of the affective flavor of expressions such as *zachtjes*. In the spirit of what was said about “genitival Case” in section 2, the bound morpheme -s is analyzed as an affixal manifestation of the categorial head n°. The appearance of dummy -s has a last resort flavor: it must appear to make the nominal status of nP recoverable at the surface.6

Let’s return to expressions such as *van je ras ras ras* and *van je hela hola*. I have just argued that *je* is the diminutive morpheme and, in section 2, I proposed that *van* is a dummy element; more specifically, it is a surface manifestation of the categorial head n°. These analyses of *je* and *van* bring me to the following analysis of the pattern *van je ras ras ras* (and other instantiations of this pattern):

(17)  

a. base structure:  
\[
[\text{ClasP} \cdot \text{je} \left[ \text{nP} \ n° \left[ \text{ConjP} \sqrt{\text{ras}} \left[ \text{Conj} \cdot \text{Conj} \cdot \cdot \cdot \cdot \sqrt{\text{van}} \cdot \text{je} \right] \right] \right] \]
\]

b. derived structure:  
\[
[\text{ClasP} \left[ n° (= \text{van}) + \text{je} \right] \left[ \text{nP} \ n° \left[ \text{ConjP} \sqrt{\text{ras}} \left[ \text{Conj} \cdot \text{Conj} \cdot \cdot \cdot \cdot \sqrt{\text{van}} \cdot \text{je} \right] \right] \right] \]
\]

As indicated in (17), the linear order *van je* results from head movement and adjunction of n° to the classifier head *je*. I tentatively propose that this movement is triggered by the affixal status of the diminutive morpheme *je*; that is, *je* must have a host to which it can be attached, quite analogously with the fact that the dummy verb to do provides a host for the Tense and inflectional features associated with the functional head T (cf. Lasnik 2000: 123) Stranded Affix Filter: “A stranded affix is no good”).

6 If *je* in (16) is a classifier that must be followed by -s, one would expect there to be other classifiers in Dutch displaying the same behavior. A plausible candidate is *stuk*(-s) (‘piece-s’) in an utterance like (i) B. Just as in *zachtjes*, -s must be present in this context. I propose *stuk*(-s) has the structure in (ii).

(i)  
A: Hoeveel broeken neem je mee? B: Twee stuk*(-s)
how-many pants take you with two piece-s

A: How many pants do you take? B: Two.

(ii)  
\[
[\text{NumP} \left[ \text{twee} \left[ \text{ClasP} \text{stuk} \left[ \text{nP} n° (= -s) \left[ \sqrt{\text{P}} \sqrt{\cdot} \right] \right] \right] \right] \]
4 The most diminutive of words

In the previous section I tried to show that *je* can occur independently, in the sense that there is no overt noun to which it can attach. Occurring on its own due to the absence of an overt noun, one might characterize this “independent” diminutive as “the most diminutive of words”. A pertinent example was *zachtjes* in (15), where -s was analyzed as a dummy element that spells out the categorial node n°. The question, obviously, arises as to whether there are more linguistic expressions featuring this “most diminutive” *je*. In what follows, I present some potential candidates, but I emphasize that further research of these ill-understood patterns is definitely needed.

Consider the following pair:

(18) a. Dit is een leuk adres-*je* voor Spaanse wijn.
    this is a nice address-*DIM* for Spanish wine
b. Dit is *JE* adres voor Spaanse wijn!
    this is *je* address for Spanish wine
    ‘This is THE address for Spanish wine’ (THE pronounced with a long vowel)

In (18a), *je* is a diminutive morpheme attached to the noun *adres*. Just as *zachtjes* in (15), *adresje* has an affective flavor. It does not imply that the address is literally small (e.g., a short street name). Rather, *je* contributes evaluative or expressive meaning (*in casu* positive valence) to the noun. Thus, objects that are big (e.g. a villa or a big car) can be referred to by N+*DIM* when the diminutive carries evaluative meaning, as in: *aardig huisje!* (*nice house-*DIM*) and *leuk autotje!* (*nice car-*DIM*). In a way, the literal meaning of the diminutive (small size) is bleached and an expressive-evaluative meaning is associated with it.

Consider next the expression *JE adres* in (18b), which has the characteristic property that the element *je*, which normally is a phonologically weak element, carries accent. The meaning associated with *je* can be paraphrased as “par excellence” (see Broekhuis & den Dikken 2016: 735). Traditionally, *JE* in (18b) is analyzed as a possessive pronoun (see Haeseryn et al. 1997). Even though a paraphrase like “the address for you (= addressee)” is imaginable for *JE adres*, the question arises as to why *je* cannot be replaced by the strong possessive pro-

---

92
noun *jouw*: *Dit is jouw adres voor Spaanse wijn!*. Instead of analyzing *je* in (18a) as a possessive pronoun, I tentatively propose that *je* is a diminutive morpheme, which, in line with Wiltschko (2005), I take to be a classifier. More specifically, I take it to be the same expressive-evaluative diminutive morpheme as in (15). By using this affective diminutive *je*, the speaker indicates that the referent of *je* adres is the representative *par excellence* of the total set of addresses.

One may wonder where the *par excellence* meaning finds its origin. Given the fact that classifiers are often used to make things countable and therefore modifiable by a numeral, one might hypothesize that the *par excellence* reading finds its origin in the (hidden, i.e. silent) presence of the Dutch numeral *één* ‘one’. Thus *je* adres has the representation in (19), where *EEN* represents a silent numeral designating singularity (and uniqueness).

\[(19) \quad [\text{NumP EEN } [\text{ClasP je } [\text{nP } t^o [P \sqrt{\text{adres }]}]]]]\]

Interestingly, this *par excellence* reading is also attested in sentences like (20a), where *eentje* consists of the numeral *een* ‘one’ and the diminutive morpheme *(t)je*. The element *er* is the so-called quantitative pro-form *er*, which arguably substitutes for nP. (20b) gives the structure of *eentje*; see Barbiers (2005) for further discussion of the numeral ‘one’.

\[(20) \quad \text{a. Je bent me er eentje!} \quad \text{‘you are really something! / you are one-of-a-kind!’} \]

\[(20) \quad \text{b. Je bent me er, [NumP één [ClasP tje [nP t_i]]]}\]

Notice that just like *zachtjes* in (15) and *adresje* in (18a), *eentje* has an affective flavor: the speaker, whose “presence” is clear from the ethical dative *me*, qualifies (and evaluates) the addressee as being unique in a certain sense.

I finish this section with another construction that possibly features the “independent” diminutive *je*.8 This construction is the italicized expression *van je*

---

8 Other constructions with a *par excellence* reading that possibly feature diminutive (i.e., classifier) *je* are *je dát* in (i) and *jé van hét* in (ii); the diacritic ‘designates that these words carry accent. I leave the analysis of these constructions for future research.

\[(i) \quad \text{a. Ik vond het niet [je dát]} \quad \text{‘I found it not je that} \quad \text{‘I wasn’t very enthusiastic about it.’} \]
welste in (21), which acts as a modifier of *krijste*.

(21) De baby *krijste* van je *wel-st-e*.

the baby screamed of *je considerable-sup-e*

‘The baby screamed enormously.’

Observe that, at the surface, *van je welste* looks a bit like the expression *van je XP*, discussed in section 3; see also *van je ras ras ras* in example (3). In both constructions, the sequence van je occurs at the beginning of a phrase. In the spirit of the analysis given in (19) for *van je ras ras ras*, I tentatively propose that *van je welste* has the base structure in (22a) and the derived structure given in (22b):

(22) a. base structure:

\[
\text{ClasP} \cdot \text{je} \ [	ext{nP} \ n^\circ \ [\text{FP} \ \text{welste} \ [\text{F} \ F \ [	ext{nP} \ n^\circ \ [\sqrt{\text{WAY}}]]]]]
\]

b. derived structure:

\[
\text{ClasP} \ [
\text{n}^\circ \ (= \text{van}) + \text{je}] \ [\text{FP} \ \text{welste} \ [\text{F} \ F \ [	ext{nP} \ n^\circ \ [\sqrt{\text{WAY}}]]]]
\]

According to this analysis, *welste* is an attributive superlative AP that modifies a silent manner noun (*WAY*). The element *van* is analyzed as the surface manifestation of the categorial node *n*\(^\circ\). I assume that *van* surfaces due to the silence of the Root; compare *do*-support in VP-ellipsis environments. The categorial node *n*\(^\circ\) raises across the attributive AP and adjoins to the classifier *je*. This yields the amalgam \([n^\circ \ (= \text{van}) + \text{je}]\).

b. *Ik vond* het niet [jé van hét]

I found it not je of it

‘I wasn’t very enthusiastic about it.’

Observe that besides *je dat* in (ia), we also find the expression *een datje* (*a that-DIM*) in fixed expressions such as (iia), (iib). In these examples, diminutive *je* is attached to the demonstrative pronoun.

(ii) a. *Hij heeft altijd wel een ditje* of *een datje*.

he has always PRT a this-DIM or a that-DIM

‘He has always something critical to say about it.’

b. *Zij spraken over ditjes en datjes*.

they spoke about this-DIM-s and that-DIM-s

‘They made small talk.’
5 Conclusion

Haegeman’s (1990, 2007) study of subject omission in English finite clauses in certain written registers, draws our attention to the relevance of peripheral data for linguistic theorizing. In this article I examined the grammatical behavior of a peripheral construction (van je XP) found in Dutch verbal lore. I proposed an analysis according to which van is a surface manifestation of the categorial node n° and je a diminutive, which was analyzed as a classifier head. I hope to have shown that, even though the van je XP pattern looks peripheral and “exotic” from the outside, the atoms and rules that underlie this construction are those that are used for the formation of simple and more familiar constructions; see also Chomsky (2015).

References


De kunst van de behoedzaamheid
The art of precaution

Karen De Clercq*

Er zijn weinig mensen die zich dieper in het moeras van Taal durven wagen dan Liliane. Er zijn ook weinig mensen die dat zo behoedzaam doen, zo beheerst en met een nooit aflatende trefzekerheid, zich volkomen bewust van alle gevaar dat op de loer ligt. Als daar zijn: onvermijdelijke focusgewassen, negatieve laag-hangende takken, linksperifere stenen op het pad en onheilsprekkende expletieven in het soppende water. Liliane draait er haar hand niet voor om. Ze weet waar de gevaren loeren en onverschrokken schrijdt ze verder, recht op haar doel af, terwijl ze haar volgers als een bedreven gids subtiel waarschuwt voor uitstekende wortels op het pad.


Gewapend met grondige kennis en analytische vaardigheden baant ze zich een weg door het dichte veen en zoekt ze alternatieve routes naar meer kennis van Taal. Op haar tocht werpt ze polders op om bedreigd land te beschermen en ontdekt ze kreken met blauwgevleugelde libellen en witte waterlelies, verborgen tussen het riet. Haar expedities monden uit in gedetailleerde kaarten die onvindbaar zijn op Google maps.

Ik wil je danken, Liliane. Om me te gidsen op mijn weg door het moeras. Om me de vrijheid te geven te duiken naar de miniscule negatiemossen onder de mattenbies. Om me te vergezellen bij dat duiken. Om me op te vissen. Om

*English translation below
me watertorkruid te tonen, gele plomp en andere vegetatieve rijkdom aan de waterkant.

Morgen trek je nieuwe laarzen aan, Liliane, begeef je je op een nieuw pad, maar er is nog veel te ontdekken en in kaart te brengen. Ik blijf je reisverhalen volgen. Laat maar horen als je er nog eens samen op uit wilt. Mijn expeditiemateriaal staat alvast klaar.

[English translation]

Few people dare to venture deeper into the swamp of Language than Liliane. Few people do this so cautiously, with such mastery and unfailing accuracy, fully aware of all dangers lying ahead. As there are: inevitable focus crops, negative low hanging branches, left peripheral stones on the path and ominous expletives in the sopping water. Liliane can handle them. She knows where the dangers lurk and bravely strides on, straight towards her goal, while subtly warning her followers, as a skilled guide, for protruding roots on the path.

The kind of precaution that Liliane displays in her linguistic work has been considered a virtue since early antiquity. And rightly so. It makes Liliane's work delicate and detailed, anticipatory and predictive. Hip-looking compasses or apps to facilitate the journey through the marshlands cannot seduce Liliane. Her approach is therefore never minimalist nor superficial. Liliane does not sail unknowingly to America like Columbus, but searches actively for the source of the Nile like Livingstone.

Armed with thorough knowledge and analytical skills, she makes her way through the dense peat bog and looks for alternative routes to a better knowledge of Language. On her journey she throws up polders to protect endangered land and discovers creeks with blue-winged dragonflies and white water lilies, hidden among the reeds. Her expeditions lead to detailed maps, untraceable on Google maps.

I want to thank you, Liliane. For guiding me on my way through the swamp. For giving me the freedom to dive to minuscule mosses of negativity under the lakeshore bulrush. For joining me in diving. For picking me up when I tripped. For showing me fine leaf water dropwort, yellow water-lilies and other riches on the waterfront.

Tomorrow you'll put on new boots, Liliane, you'll be walking a new path, but there's still a lot to discover and map out. I will continue to follow your travel stories and maps. Let me know if you want to go out exploring together again. My expedition gear is already packed up.
‘What would you rather me say?’

Marcel Den Dikken

Abstract
In varieties of English, the combination of would and rather (also: sooner/as soon/as well) can be followed not just by a bare infinitive (as in they would rather leave) or by a finite clause (they would rather (that) I (would) leave), but also by an infinitive with an accusative subject (as in they would rather me leave), which can even be coreferent with the matrix subject (I would rather me leave). This short paper focuses on this Acl-infinitival construction. It shows that the infinitival clause is a fully clausal complement of rather, capable of harbouring sentential negation and constituting a local binding domain for its subject, whose accusative is not an assigned case. The paper closes on some remarks about the evolution of this construction, against the background of the form and distribution of the subjunctive.

Sentences of the type in (1) feature a degree-modified dispositional adjective or adverb (rather, sooner, as soon, as well) followed by a bare infinitival clause with an accusative subject.

(1) a. they would rather me leave
   b. they would sooner me leave
   c. they would just as soon me leave
   d. they would just as well me leave

Though this construction seems particularly common with me (representing the speaker) as the accusative subject of the infinitive, it is not restricted to me, as witness sentences such as I would rather him/her/us/them be happy. Though the initial exemplifications in [1] are made-up sentences to keep them simple and directly comparable, the bulk of the examples in this paper are attested sentences culled from the internet, and checked with native-speaker linguists. In what follows, exemplification will generally be confined to rather.
The infinitival clause is transparent for extraction of its object (as in (2a)), its subject (as in (2b)), and even a modifier (as in (2c)).

(2)  
   a. what would you rather me say?  
   b. who would you rather kiss you like that?  
   c. how would you rather me do this?

The adjective cannot be *wh*-fronted, however, despite the fact that it can be intensified with *much*:

(3)  
   a. you would much rather me do this  
   b. *how much rather would you me do this?

In precluding extraction of the adjective, the construction in (1) patterns like the one in (4), and unlike that in (5).

(4)  
   a. you would much rather that I do this  
   b. *how much rather would you that I do this?

(5)  
   a. you would much rather do this  
   b. how much rather would you do this?

In (5a), *much rather* is a modifier of the projection of *do*, which is the main verb of a single clause. In (4a), we are evidently dealing with a biclausal construction, with the *that*-clause serving as the complement of *rather*. This straightforwardly explains the contrast between (4b) and (5a) in the latter but not in the former, *(how) much rather* is a constituent. The fact that (3b) behaves the same way as (4b) suggests that in (3a) and (1), too, what follows *rather* is a clausal complement:

(6)  
   \[ \text{AP A=rather [CLAUSE]} \]

The clausality of what follows *rather* is perfectly apparent for variants of (1) in which the infinitive is adorned with the infinitival marker *to*, as in (7a), which, like (1), allows extraction, as shown in (7b).

(7)  
   a. you would rather me to do this  
   b. what would you rather me to do?

---

2 Again, the pattern is not limited to *me*: sentences such as *what would you rather him say?* (directly parallel to (2a)) and *what would you rather him be to you?* (with extraction of the predicate of a copular infinitive) occur frequently as well.
But for the bare infinitive in (1) as well, there are clear indications that it is a clausal constituent. Thus, sentential negation is grammatical in the bare-infinitival clause, as shown by the fact that not in (8) licenses negative polarity items:

(8) you would rather/sooner/just as soon/just as well me not do anything

Since sentential negation is by consensus (see Zanuttini 1997, references there, and work in its wake) dependent on T, the bare infinitive in (8) must at a minimum be a TP – i.e., a full clause.

The hypothesis that the bare-infinitival constituent in (1) is a full clause helps us explain the otherwise quite intractable fact that the subject of the bare infinitive can be a pronoun coreferential with the matrix subject:

(9) a. I would rather me leave
    b. I would rather me die than you

For (9b), the acceptability of me could perhaps be ascribed to the fact that this pronoun is a contrastive focus (contrasted with you). But in (9a) the subject of the infinitive is not contrastive. The fact that it allows itself to be coreferential with the matrix subject indicates that the bare-infinitival constituent is a local domain for binding, just as in I would rather that I leave. The postulation of clausal structure for the bare infinitive contributes to making this understandable.

By itself, however, the clausal (i.e., TP) status of the complement of rather does not immediately facilitate a coreference relation between the matrix and embedded subjects: for bare Acl-infinitives embedded under causative or perception verbs, the grammaticality of clausal negation (as in (10a)) suggests full clausality; yet coreference of the pronominal subject of the infinitive with the matrix subject is impossible, as (10b) shows.

(10) a. I made/saw him not eat anything
    b. *I made/saw me leave

The difference between (1) and causative and perception verb constructions lies in the category of the selector of the bare Acl-infinitive: an element of category A in the former and a verb in the latter. Thanks to the fact that in causative and perception verb constructions, the infinitive’s selector is of the same category as the head of the infinitival complement, the two domains engage in what is

3Like (1) and (2), the pattern in (9) is attested with accusative pronouns other than me, as in he would rather him be dead or they would rather them do work around the house.
variably called ‘reanalysis’, ‘restructuring’ or ‘clause union’. In (1), where the categories of the bare infinitive and its selector are different, such integration is impossible. The infinitival TP in (1) thus remains an independent binding domain, making (9b) grammatical.

Though opaque for binding, the fact that the infinitival clause in the complement of rather is a selected constituent renders it transparent to extraction: we saw this in (2) for disjoint reference cases; (11) shows that in coreference contexts, though much rarer, wh-extraction is also possible — even for a predicate nominal (what in (11) is a predicate), which strongly resists extraction from islands (*what don’t you know whether to call these linguists? vs. *which linguists don’t you know whether to call ‘generativists’?).

(11) what would they rather them be named?

With respect to extraction, (1) and (9) are similar to likely+infinitive constructions, in which the infinitival complement to the adjective likely is likewise transparent to argument and non-argument wh-extraction (what is he likely to say?, how is he likely to solve the problem?).

What could be the source of the accusative case of the subject of the AcI-infinitive? For causative and perception verb constructions, the answer is straightforward: the matrix clause contributes an accusative case feature, assignable to the subject of the infinitival clause. But in the construction in (1), the selector of the infinitival clause is of category A, incapable of assigning structural case. The morphological accusative case of the subject of the infinitival clause in (1) can be dealt with in either of two ways. One would be to treat it as a manifestation of default case, which in English is indeed accusative. Alternatively, it may be possible to invoke Marantz’s (1991) notion of dependent case: nominative case in the matrix domain is already used up by the matrix subject, so the subject of the infinitival clause gets the dependent accusative instead. The feasibility of a dependent-case approach to the accusative in (1) will depend crucially on whether they and me in these sentences belong to the same local domain. Above, we saw that the infinitival clause embedded under rather is an opaque domain for binding; yet at the same time it is a transparent domain for wh-extraction. If the infinitive’s complement status is sufficient to include it in same local domain as the matrix subject for the purposes of dependent case assignment, it will be possible to get

4Though I’d rather me than you occurs as a complete utterance, it is arguably always elliptical: the syntax features a clause with a pronominal subject rather than just a pronoun in the complement of rather.
dependent accusative case assigned to the subject of the infinitive. In view of the fact that passivisation (with concomitant promotion of the structural accusative to subject) is independently impossible in the rather+infinitive construction (because would rather does not passivise), it is difficult to ascertain whether the case of the subject of the infinitive is structural (i.e., dependent) or default.

In closing, I would like to make a few speculative remarks about the evolution and spread of the construction illustrated in (1). It seems to me likely that pairs of sentences such as the following play a major role in the development of the rather+ACC+INF construction:

(12) a. you would rather (that) I were more serious
    b. you would rather (that) I be more serious

In (12a), we are dealing with a subjunctive subordinate clause, whence the nominative subject, I. On the surface (and perhaps also in a deeper sense), the subjunctive form of the English verb is indistinct from the bare infinitive. So for the version of (12b) lacking the complementizer that, the complement clause is easily reanalysed as an infinitival clause. Such a reanalysis deprives the subject of that clause of its nominative case, and leads to a (default/dependent) accusative case form, as in you would rather me be more serious – an instantiation of the pattern in (1).

Though the negation and pronominal coreference facts reviewed above have led me to conclude that the bare infinitive embedded under rather is fully clausal, a logical next step in the development of the construction type would be for the infinitival constituent to be analysed as a small clause. Once this happens, we expect to be able to find bare non-verbal predication structures with an accusative subject in the complement of would rather. Indeed, this seems to have become a reality, judging from the occurrence of sentences such as the ones in (13a)–(13c) (with disjoint reference, à la (1)) and (13d) (with coreference of the matrix and embedded subjects, as in (9)):

(13) a. AirTran would rather me stinky
    b. he would rather them dead
    c. I would rather him sick now than when he is in school
    d. I would rather me sick than you

The previous paragraphs paint a preliminary picture of a construction type which, as far as I am aware, has not received detailed attention in the theoretical literature to date. It goes without saying that much more could and should be said.
about *what would you rather me say?* and its ilk. I hope that the honouree of this webschrift will say that what these initial notes say is well-said. But perhaps she would rather me say just simply: Thank you very much, Liliane, for all the wonderful linguist(ic)s that you have given the world, for the innumerable ways in which you have contributed, empirically as well as theoretically, to the generative enterprise, and for all the great fun we’ve had.

**References**


In contrast to most of the other contributors to this webschrift, I have not known Liliane Haegeman in any of the usual professional roles: she has never been my department colleague, teacher or mentor (in the obvious sense); indeed, over the years we have met in person no more than a handful of times. Yet for most of my career Liliane has been a constant familiar, guiding and supporting my efforts to understand syntax and language acquisition better than would otherwise have been possible. A ‘shoulder angel’ of sorts—always of the good kind. Why she ever took on this role is not something I can explain, though I suspect I’m not the only one to benefit from her altruistic advice and generosity of spirit. In gratitude, I offer the following recollection of two early occasions where her interventions were most telling. As it turns out, neither of the questions that I was trying to address received a satisfactory solution at the time. This most likely reflects failures on my part, and since these shortcomings would have gone unnoticed but for her, perhaps I shouldn’t be so grateful. On the other hand—and this is the justification for this contribution—the two problems have remained something of an embarrassment for standard versions of generative theory ever since, and it seems right that they should be dusted off in her honour. With luck, this will encourage a new generation of linguists to find more imaginative solutions, or at least not reinvent the wheel.

1 Subjects, Case and Negation in Hiberno-English

Almost exactly thirty years ago, I wrote my first term paper in graduate school. Titled *A Case for Default Values*, it examined the origins of certain constructions found in Hiberno-English (HE), first brought to general attention by P. L. Henry (1977).¹ In the original paper, subsequently presented at WECOL (Duffield 1989),

¹P. L. Henry is no relation of Alison Henry, as far as I’m aware. In his work, Henry (1977) uses the term Anglo-Irish, which is now generally dispreferred: see Filppula (2002), for discussion. I
the main concern was to account for the acceptability of HE examples such as those in (1) and (2) below, all of which are grammatically unacceptable in standard varieties of English (on both sides of the Atlantic):

(1)  
   a. ‘T’is an aise to the gate, they to be married.’
   b. ‘T’was a loss to the country, Michael to die.’
   c. ‘She to go and he to follow her, t’is the best anymore.’
   d. ‘Typical,’ said Morrissey, as she walked away, ‘a female like that to spoil a funeral.’

(2)  
   a. ‘...surely not knowing the way, they just give him a slap or something like that and he, oh he to be afraid of the life of him.’
   b. ‘D’you mind the day, and we in the old castle?’
   c. ‘I heard the hens cacklin’, and I went over to see what it was, and here it was a fox, and he with a hen.’

As should be clear, the theoretical interest of such sentences lies in the nominative case-marking on non-finite subjects, most evident in those examples involving pronominal subjects (1a), (1c): then and still, these sentences offer a prima facie challenge to traditional generativist assumptions linking nominative case to finiteness (Chomsky 1981, Pesetsky & Torrego 2002; cf. Szabolcsi 2009, esp. McFadden & Sundaresan 2011).

The specific question I was concerned with at the time was whether the apparent ‘default subject case’ option had arisen in Hiberno-English through contact with Modern Irish; alternatively, whether the Hiberno-English infinitival construction was the vestige of a licensing option previously observed in ‘mainland’ Early Modern English. Both hypotheses are supported by circumstantial evidence. On the one hand, as discussed in Chung & McCloskey (1987), Modern Irish freely allows overt subjects in small clauses in the absence of any obvi-
ous licensor—albeit these are marked with accusative, rather than nominative, case—see [3] below; on the other hand, Late Middle English and Early Modern English apparently tolerated nominative subjects in the same non-finite contexts, as evidenced by the examples in (4) from Visser (1963), see also Sundaresan & McFadden (2009).

(3) a. Bhual mé leis, agus [s é dhá mhíle as baile].
   struck I with-him and him.ACC 2 miles from home
   ‘I met him, as he was two miles from home.’

b. Tháinig sé isteach, agus [s é iontach sásta leis féin].
   came he inside, and him.ACC very happy with-him self
   ‘He came in very happy with himself.’ (Chung & McCloskey 1987)

c. Is mór an suaimhneas don gheata [s iad a bheith]
   is great the ease to-the gate them-ACC to be-VN
   married.
   ‘Tis an aise to the gate they to be married.’ (Henry 1957; cf. (1a) above)

(4) a. I to make me blith or glad ... [th]at nu mai be... [14th C]

b. A king to kepe his lygis in justice, Without doute that is his offise...
   [1385]

c. Men to seye to women wel, it is best, And nor for to despise hem ne
drape.
   [1402]

d. A preest for to freli take and chose of alle maidens to hem a wijf...was
   allowed of Poul
   [1449]

e. Thou to love that loueth not the, is but grete foly.
   [1470]

f. She to dy so dangerously ... that was the thing that greued me so.
   [1570]

Partly due to the paucity of historical evidence, the provenance issue was not resolved, and this parochial phenomenon might have completely escaped further attention, were it not for the availability of another non-standard subject type found in (some of) the same HE varieties, namely, NPI-subjects in finite main clauses, as shown in (5):

(5) a. But from that day out, anyone that was on the meitheal, or anyone in the parish never said a bit t’him, or never done...made a move to have sport on him...
b. Aha. Although anybody don’t seem to like to live in Russia...They’re all trying to get out of it...
c. you couldn’t pick a daisy but it was a sin. Now, anything is no sin...

In a follow-up conference paper, I sought to relate the grammaticality of these two constructions to a third HE anomaly, namely, the phenomenon of Singular Concord (SC) observed in Belfast English, discussed in A. M. Henry (1992, 1995), whereby plural lexical NPs in canonical position can appear with default (singular) agreement—in contrast to pronominal and/or inverted subjects, which always require plural agreement. Compare the examples in (6):

(6)  a. Them eggs is cracked/*They is cracked/They are cracked.
    b. Them ones wants more than them other ones/What do/*does them ones want?

Extending Henry’s (1992, 1995) account of SC, it was proposed that matrix clauses in HE and Early Modern English varieties differ(ed) from Standard varieties in containing a higher functional projection (AgrP) above TP (AgrP\NegP^TP), and in allowing default case to be assigned in (non-root) Spec, TP; cf. McCloskey (2001). Given this parametric option, underspecified DPs lacking person features (6: them eggs) need raise no higher than TP for Case, and need not agree (is shows default singular agreement). Subject NPIs (5: anyone was...) are then licensed in virtue of having a case-marked copy within the scope of negation (Duffield 1992/1993). On this approach, the diachronic loss of nominal subjects in infinitivals in Standard English—as well as absence of the other two properties—was taken to stem from the loss of AgrP, and subsequent re-analysis of the root projection as TP, as diagrammed in (7); cf. Roberts (1993).

(7)  [AgrP [NegP [TP [NOM] ... ]] \rightarrow [TP [NOM] [Neg [ ... ]]]]

The analysis is undoubtedly outdated, most notably in its construal of AgrP as an autonomous functional projection, and in its appeal to Case as a licensor and as principal driver of syntactic movement to Spec, TP (Chomsky 2001, cf. Nevins 2004). Even at the time, the analysis was probably seriously flawed. Yet it did at least draw a connection between two kinds of anomalous subject that most generative researchers had totally ignored up to that point. Just as importantly from a personal point of view, the analysis apparently had sufficient merit to persuade Liliane to include it in a special issue of Rivista di Linguistica, on the syntax of sentential negation (Haegeman 1993). Which in turn gave me my first journal
article and—I suspect (though cannot prove)—led to my first job after graduate school.

2 Pro-drop in Early Child German

If Liliane’s invitation helped me into my first paid employment, as a post-doc in Harald Clahsen’s LEXLERN project on child German, her support of the research begun there also helped me into the next position, at McGill. Once again, the focus of this work was on anomalous subjects: this time, however, on null-subjects in German.

As is well-known—implicitly to native-speakers, explicitly to professional linguists and German language teachers—Standard German is not a pro-drop language: as in English, the subjects of finite clauses must be pronounced. Yet viewed from the bottom up—which is to say, from the PLD perspective of the child learner—this ban on null-subjects is quite unexpected. This is for three reasons that also distinguish German from English: first, German is a ‘topic-drop’ language, allowing omission of initial topics (8); second, it allows null-expletives (except in initial position) (9); finally, it has a relatively rich person-number verb-agreement paradigm, something that is often considered key to pro-identification; Rohrbacher (1993), cf. Bobaljik (1997).

(8) a. (Ich) habe es gestern gekauft.
   (I) have it yesterday bought
   ‘I bought it yesterday.’

b. (Er) sagt, daß *(er) es gestern gekauft hat.
   (he) says, COMP (he) it yesterday bought has
   ‘He says he bought it yesterday.’

c. Gestern habe *(ich) es gekauft.
   yesterday have (I) it bought
   ‘I bought it yesterday.’

4 As hard as it is to believe now, having one journal article published before completing graduate school counted as a distinguishing achievement in the early 1990s.

5 Or it might have been my next job (University of Sheffield). The lag between initial presentation of the data at acquisition conferences in 1992-93, submission of the initial draft of the paper, hiatuses between revisions, and final publication in 2008, spanned 15 years, during which time Liliane’s support both as known commentator and as anonymous reviewer proved invaluable.
(9) a. ... daß ec dem Mann das Buch geschenkt wurde.  
    that the. DAT man the-NOM book presented  
    [Biberauer 2008]  
    became  
    ‘... that the book was given to the man (as a present).

b. ... weil ec ja doch Linguisten Kammermusik spielen  
    ‘... since there are linguists playing chamber music.’

So what do young German children think about null subjects? That was the focus of my two-year study in Düsseldorf, which compared the Strong Continuity/Very Early Parameter Setting claims of Rizzi (1994, 2000), Wexler (1994, 1998) and others, with the Weak Continuity/Structure Building approach of the LEXLERN project (Clahsen & Penke 1992, Clahsen et al. 1994, 1996); cf. Haege- 
man (1996). The two approaches to early syntax diverged most sharply on the issue of postverbal null-subjects in finite V2 clauses, termed ‘Rogues’ in Duffield (1992/1993, 2008): whereas Strong Continuity approaches predicted their non-occurrence—Root Infinitive subjects aside, German children should never entertain the possibility of a null-subject grammar—the Weak Continuity predicted the opposite, namely, it was expected that children should pass through a stage after the development of an underspecified AgrP (but prior to the development of CP) during which Rogues should be a grammatical option. Data from four out of six LEXLERN 6 children whose transcribed corpus data I analyzed provide rather clear support for the latter approach: not only did all of these children produce Rogues, such as those illustrated in (10), but in each case, these were only observed at significant levels within one developmental window, just before their production of finite complement clauses.

(10) a. Einen Strohhalm mach’ ec jetzt. (Matthias.22: 089)  
    a-ACC straw make-1SG now  
    ‘I’m making a straw now.’

b. So sieht ec das nicht. (Katrin.07: 021)  
    so see-3SG that not  
    ‘That way, he doesn’t see it.’

---

6In fact, the study drew on data from three separate corpora, to which the project had access.
c. Die Margot hab’ ec gehört. (Svenja.06: 134a)
   the Margot have-1SG heard
   ‘I have heard Margot.’

Once again, my interpretation may be flawed—the framework has surely been superseded, AgrP is so 1990s—but the data remain, as a thorn in the side of those who reject parametric approaches, and who discount any significant learning from the input. And once more, had it not been for Liliane’s support and encouragement, these data would likely never have seen the light of (theoretical) day. Even with her help, the journey took fifteen years!

Thus, I use this opportunity to thank Liliane for supporting research on things that shouldn’t be present, but are, or should be absent, but aren’t, for being so. While we are on the subject.

References


Haegeman, L. 1996. Verb second, the split CP and initial null subjects in Dutch. *GenGenP* 4. 93–120.


Some thoughts on the early Dutch null subject

Maria Teresa Guasti

Between the age of 2;0-3;6, children speaking non-pro drop languages (English, Flemish, French, Danish) omit the sentential subject from their sentences. This occurs more often if the verb is nonfinite, but it occurs with average percentages that go from about 50\% to 20\% in finite sentences as well. For example, a French-speaking child could utter (1).

\[(1)\] est trop gros
(he) is too big
(Philippe, 2;2)

Italian-speaking children also omit the sentential subject, but this is not surprising since their grammar allows it. Dutch-speaking children omit the sentential subject. Although this should not surprise, as topic drop is licit in the target language, Haegeman (1995) has challenged the claim that children’s omissions were target-consistent and has convincingly shown that these omissions had to be put in the same basket as subject omissions in English, French. She has claimed that the early null subject of Dutch-speaking children is a root phenomenon. She has shown that omission occurs from matrix, but not from embedded clauses, that is, subject omission is limited to declarative sentences and non-existent in wh-questions. However, these properties characterize Dutch topic drop as well. To support her claim, Haegeman has provided data showing that subject omission in early Dutch decreases between 2;0 and 3;0 years and this decrease parallels that of root infinitives, another phenomenon observed in the early grammar of children speaking a wide varieties of languages. On this basis, she argues that the early Dutch null subject is a developmental phenomenon that fades away as children’s grow older and their cognitive resources increase. Specifically, Haegeman, following Rizzi (1994), has proposed that Dutch-speaking children were applying clausal truncation to their sentences and produced null
subjects in the Spec of the root. She has also shown that this phenomenon is not only present in child Dutch, but in diaries written by several well renowned adult writers (Haegeman 1990, Weir 2012), although adults’ null subjects may be different from children’s ones. Thus, in omitting the root subject children were exercising an option available in Universal Grammar. Of course in the case of writers, it is not short of cognitive resources that is behind the phenomenon, but something else.

The phenomenon of subject omission has been extensively investigated since Hyams (1986), but mostly in non-null subject or non-drop languages, with some exceptions. When the productions of children exposed to null subject or topic drop languages have been examined, it were to show that subject omission by these children were qualitatively and quantitatively different from those of children exposed to a language where omission was not target consistent. One example is Wang et al. (1992). Mandarin is a topic drop language, whereby both subjects and objects can be dropped under appropriate contextual conditions. Wang et al. (1992) have shown that Mandarin speaking children drop subjects and objects 46% and 22% of the time, respectively, at 3;0 years of age. Their English-speaking peers do so much less, especially for objects, i.e., 33% and 4%.

It is clear that these figures tell us that the null subject phenomenon in child Mandarin and in child English emanates from different grammatical sources. An important and disregarded finding from Wang et al. (1992)’s paper is that Mandarin-speaking children omit in a similar way as adults, as far as the asymmetry subject vs. object is concerned. Adults drop subjects more often than objects, i.e., 36% and 10%, respectively. However, looking at Wang et al. (1992)’s data, we can observe that during children’s development there is a decrease in null subjects: subjects are omitted 56% around age 2;0, 46% around age 3;0 and 38% around age 4. In other words, it is around age 4;0 that children’s subject drop and adult subject drop quantitatively matches, as already observed in Guasti (2017). Wang et al. (1992) did not look systematically at the pragmatic conditions of subject drop. They merely say that “the null subject was sometimes clearly related to an antecedent… In other cases, the referent of the null subject was not previously mentioned in the discourse, although it was usually understandable from context” (Wang et al. 1992: 233). Therefore, we can infer that most of the time the use was appropriate, but a more systematic analysis would have been appropriate.

A developmental decrease in subject omission is also observed in early null subject languages. Valian (1991) reports that Italian adults omit the sentential subject 64% of the time. Serratrice (2005) shows that, across four MLU stages,
omission of the sentential subject decreases from 80% to 65%. At the same time, she also demonstrates that Italian-speaking children from their earliest multiword productions (around age 2;0) are sensitive to the pragmatic constraints regulating subject omission (e.g., informativeness). They omit more 1st and 2nd person subjects than 3rd person subject, as the former are always recoverable from context. They use overt subjects when it is needed because the antecedent is not easily recoverable. However, we when look at the quantitative data she provides, we realize that there is a certain amount of null subjects whose use is not accounted for by her pragmatic constraints and in fact she recognizes that “...although children are sensitive to discourse-pragmatic features in their choice of referential expressions, this sensitivity becomes more fine-tuned over time” (Serratrice 2005: 457).

These findings invite us to think that subject omission in early Dutch (and in early Italian as well) conceals two phenomena: the phenomenon found in other early non-topic drop or non-null subject languages and the adult topic drop option, as proposed in Guasti (2017). This would reconcile Hageman’s claim that Dutch-speaking children are using the option exploited by English-speaking children with the claim by de Haan & Tuijnman 1988 and Verrips & Weissenborn 1993 that children are using the adult topic drop option. To further explore this hypothesis, one would also need to look more closely at the features of topic drop in spoken Dutch. Trift (2003), through a judgement experiment with Dutch-speaking adults, has shown that dropping 1st person subjects is more tolerated than dropping 2nd and 3rd person subject. One may wonder if this is the case in early Dutch. Some other similar insights come from work by Frazier. Based on adult data, Frazier (2015) has proposed a view that is different from that taken here. According to her, subject omission is not a grammatical option, as assumed in Haegeman (1995), but it is due to a performance error that occurs when the subject is highly predictable. She bases her claim on an experiment carried out by Mack et al. (2012) combined with a second experiment she carried out. In the first experiment, English-speaking adults heard mini-dialogues, in which the final sentence had a 1st or 3rd person subject, and either it was in the present or past tense. The subject was degraded, so that it was not clear whether it was pronounced or not. Adults were invited to repeat the sentence. It was found that they restored the unheard subject more often with a 3rd person than with a 1st person verb and more often in the past than in the present tense. In the second experiment, the same result was obtained with non-sense verbs, supporting the conclusion that the phenomenon is not lexically driven. Frazier interpreted these findings as evidence that adults expect a subject and restore it more often
in certain conditions than in others, specifically they restore it when it is less predictable. Frazier only used sentences with possible root null subjects. Therefore, we cannot establish whether the accepted null subjects (i.e., those not restored) emanate from a grammar that allows the root null subject or is really the expression of a speech error, as Frazier claims. It would be interesting to extend Frazier’ study to contexts with highly predictable subjects in wh-questions or embedded questions to establish whether adults restore the subject more often in these contexts than in root contexts, as our grammatical approach would suggest. It would also be interesting to see what children would do in these same contexts, if the experiment is feasible. This would give us some insights not only into their syntactic knowledge, but also into their pragmatic knowledge.

In conclusion, I have attempted to reconcile two opposing views concerning subject omission in early Dutch and suggested that Dutch-speaking children know that their grammar allows topic drop, but at the same time they exercise another option found in other grammars. Second, I have suggested a new path of investigation which is concerned with the pragmatics of null subjects, i.e., the conditions that lead speakers of non-null subject grammars to tolerate omission.

References


Negative concord and Old English clause structure: some afterthoughts

Eric Haeberli*

1 Introduction

Old English (OE) has word order properties that are reminiscent of the modern West Germanic asymmetric Verb Second (V2) languages like Dutch or German. In main clauses, the finite verb tends to occur towards the beginning of the clause whereas placement towards the end of the clause is common in subordinate clauses. To capture this similarity, van Kemenade (1988) proposes that OE can be structurally analysed like modern West Germanic. In basic X-bar theoretic terms, this means that the finite verb moves to C in main clauses while it stays in T in a head-final TP in subordinate clauses. However, it has been shown in subsequent work that there is evidence for head-initial structure in subordinate clauses and that therefore the headedness of TP may be variable in OE. Different types of evidence have been put forward to support the hypothesis that the verb can occur in a head-initial projection in subordinate clauses, with the main data being related to the distribution of particles, pronouns, stranded prepositions

*I am very happy to dedicate a second paper of mine to Liliane Haegeman. A single one would simply not have been enough for someone special like Liliane. Here, I would like to express my particular gratitude to her for getting me started in linguistics. Making use of further data and recent theoretical proposals, the present trip down memory lane provides a short addendum to our joint work, which marked the beginning of the gratifying privilege I had of being Liliane's PhD student.

1 For simplicity's sake, I will refer to structures here that include head-final projections. However, this choice is not likely to substantially affect the main points made in this paper. My main focus will be an empirical generalization on OE and West Flemish syntax, and, as far as I can tell, there are no obvious advantages or disadvantages in capturing this generalization in terms of an approach that includes head-final structure and one that bans them (cf. Haeberli & Haegeman 1995: 103-107 for some discussion).
and negative objects (Pintzuk 1993, 1999, 2005) and to Negative Concord readings (Haeberli & Haegeman 1999). In this paper, I will reconsider the evidence from Negative Concord by adding some quantitative data and by examining relevant word orders in light of the most recent proposals that have been made with respect to the structural analysis of OE.

2 OE subordinate clause syntax

In OE subordinate clauses, the finite verb frequently occurs at the end of the clause. This is shown in (1).

(1)  swilce he wið his dohtor sume digle spæce sprekan wolde
     as he with his daughter some secret speech speak would
     ‘as he would speak some secret speech with this daughter’ (coapollo,ApT:1.10.10)

In (1), the finite auxiliary is preceded by the subject pronoun, a PP, an object, and a non-finite main verb. Both clauses can be analysed by assuming that the subject is in SpecTP, the finite verbal element in a head-final T and the other constituent(s) in the complement of T. As van Kemenade (1988) already points out, however, the finite verb is not always in final position in OE subordinate clauses. This is illustrated in (2).

(2) a. þæt ic mihte [God forbeodan]
     that I could God forbid
     ‘that I could forbid God’
     (coaelive,+ALS[Peter’s_Chair]:186.2398)

b. ðæt se recere ða ðeawas & ða unðeawas cunne [wel
     that the teacher the virtues & the vices can well
tosadan]
     distinguish
     ‘that the teacher can distinguish virtues and vices well’
     (cocura,CP:20.149.16.1019)

In (2), the finite auxiliary is followed by the non-finite main verb and its complement (2a) or an adjunct (2b). Given cases like those in (2), van Kemenade (1988) proposes a range of rightward movements for OE that derive word orders in which the finite verb is not at the end of the clause. For example the
subordinate clauses in (2) can be argued to involve movement of the bracketed constituent from the left of the auxiliary to the right, a process that has been referred to as Verb Projection Raising (VPR) in the literature on varieties of modern West Germanic.

If, in (2b), the subject is in SpecTP and the finite auxiliary in T, an analysis in terms of VPR seems to be inevitable. The word order in (2a), however, could be derived either through VPR or simply through a head-initial TP structure in which T selects a head-final VP to its right. Strong evidence for the availability of the latter option, i.e. head-initial TP, in OE is provided by Pintzuk (1993, 1999, 2005). Her approach is to look at clauses that must clearly be head-final and then to examine what types of elements can and cannot undergo rightward movement in such clauses. Under the assumption that with a head-initial TP only one XP position is available before the verb in T (i.e. SpecTP), Pintzuk considers those subordinate clauses as unambiguously head-final in which at least two heavy constituents precede the finite verb. Focusing only on such clauses, Pintzuk identifies several elements that can generally not occur to the right of the verb: particles, object pronouns, stranded prepositions and negative object. But in subordinate clauses with a single constituent before the finite verb, these elements do occur after the verb. Pintzuk therefore concludes that particles, object pronouns, stranded prepositions and negative objects are diagnostic elements for head-initial structure. Since the diagnostic elements cannot undergo rightward movement in unambiguously head-final clauses, their occurrence to the right of the verb must be the sign of a head-initial TP structure. Pintzuk therefore proposes that OE shows variation with respect to directionality: TP can be both head-initial and head-final (the double base hypothesis).

3 Negative Concord and OE clause structure

Haeberli & Haegeman (1995) (henceforth HH) provide independent cross-linguistically based evidence in favour of Pintzuk’s hypothesis that head-initial TP structure can be found in OE. HH’s argument is similar to Pintzuk’s as it is based on a phenomenon that does not seem to be possible with head-final structure, but nevertheless occurs in OE subordinate clauses. The empirical domain that HH consider is Negative Concord (NC), the phenomenon whereby two or more negative elements in a clause do not cancel each other out but together express a single negation.

HH’s starting point is an observation made by Haegeman (1995) for West
Flemish (WF). WF has an asymmetric V2 syntax with verb-final subordinate clauses comparable to Dutch and German. But in contrast to those languages, WF also has VPR and productive NC. An illustration of NC is given in (3) (from Haegeman 1995: 133):

(3) da Valère an niemand niets nie gezeid oat
    ‘that Valère had not said anything to anyone’

What is of interest for our purposes now is that a negative constituent contained within a VP that has undergone VPR cannot enter an NC relation with an element outside this VP. I will refer to this generalisation with the label *NC-VPR. It is illustrated in (4), where only the Double Negation reading is possible.

(4) dan-ze niemand nie willen niets zeggen
    ‘that they do not want to say nothing to anyone’ (Double Negation)
    ‘*that they do not want to say anything to anyone’ (NC)

As HH point out, *NC-VPR makes an interesting prediction for OE. Like WF, OE is a NC language. Sentential negation is expressed by the preverbal negative marker ne, which can co-occur with one or more other negative elements (three in (5)) to express a single negation.

(5) þe næfre nan man ne geseah ær on nanum lande
    ‘that no man has ever seen before in any land’
    (cootest, Exod: 34.10.3577)

The prediction then is the following: If *NC-VPR also holds in OE and NC is not possible with a negative element occurring inside a VP that has undergone VPR, the configuration ‘Neg1-Aux-Neg2-V’ with an NC reading should only be found with clauses that can be analysed as involving head-initial TP as the relevant word order could be obtained without VPR in such cases. As for clauses which must be analysed in terms of a head-final TP, NC readings with a negative element between the finite auxiliary and the non-finite main verb should be ungrammatical and therefore not be found in OE.

HH claim that this prediction is largely borne out. Clearly head-final subordinate clauses with VPR generally do not have a negative element in the VP moved to the right. However, HH identify eight examples that could potentially be prob-
lematic. They propose an analysis for six of them and speculate on the status of the remaining two. Overall, HH conclude that *NC-VPR also seems to hold in OE, and that the regular occurrence of the order ‘Neg1 Aux Neg2 V’ can be related to head-initial TP, confirming Pintzuk’s double base hypothesis.

In the remainder of this paper, I briefly re-evaluate HH’s conclusions for two main reasons. First, HH do not provide any quantitative evidence. This weakens their claims somewhat as a crucial part of their argument is based on the hypothesized ungrammaticality of an option in OE (NC with VPR in clearly head-final clauses). Ungrammaticalities can never be conclusively established in corpus data, but quantitative evidence is useful in that the likelihood that the absence of an option in a corpus is a sign of ungrammaticality increases with the number of examples in which this option could have occurred but did not. A second reason for reconsidering HH’s observations is that more work on OE syntax has been carried out over the last 20 years, and we may wonder what the status of the examples that HH have identified as potentially problematic is within more recent analyses of OE.

3.1 *NC-VPR: Quantitative evidence

My first goal is to provide a quantitative analysis of the interaction between NC and VPR. In contrast to HH, who used A Microfiche Concordance to Old English (Healey & Venezky 1980), I will base myself on the YCOE (Taylor et al. 2003). The main issue I will consider is the following: Under the assumption that *NC-VPR holds for OE as it does for WF, HH propose that all subordinate clauses of the type ‘Neg1-Aux-Neg2-V’ must be the result of head-initial TP. We therefore get the following prediction P that needs to be tested quantitatively:

(6) P: The sequence ‘Neg1-Aux-Neg2-V’ does not occur in clearly head-final clauses in OE.

With respect to defining “clearly head-final”, I will start with the minimal and most constrained hypothesis (but cf. section 3.2 below for some further discussion). In a head-initial structure, the finite auxiliary is under T and the subject is in SpecTP. The minimal assumption is therefore that they are adjacent. Thus, head-final structure is required if one or more constituents intervene between the subject and the finite auxiliary.

In order to test P, I have collected all OE subordinate clauses containing a finite auxiliary, a non-finite main verb and at least two negative items. In all cases,
one of the negative items is the preverbal negative marker *ne*, which can express negation on its own and is generally analysed as a prefix on the finite verb. An NC relation is then established between *ne* and one or more additional negative elements in the clause. Overall, there are 631 clauses that meet the requirements described before. However, not all of them are equally relevant for testing P. In 146 clauses, the second negative element is the subject occurring in SpecTP. In these cases, a violation of *NS-VPR* would not have been possible as the subject in SpecTP could not have been included in a VP undergoing VPR.

This leaves us with 475 clauses containing *ne* plus one or more negative non-subject XPs. But once again, not all of these could have given rise to a violation of *NS-VPR*. For that to be possible, there has to be one additional non-subject constituent which could intervene between the subject position and the auxiliary and thus indicate clear head-final structure. Focusing then on all subordinate clauses that contain a negated auxiliary, a non-finite main verb, at least one negative non-subject XP, and at least one other non-subject XP, we are left with 266 clauses that could potentially have given rise to a violation of *NC-VPR* if the four elements had been placed in the order ‘XP-*ne*Aux-NegYP-V’.

Thus, the pool of examples that allow us to test P is relatively substantial. If we now examine it for the occurrence of ‘XP-*ne*Aux-NegYP-V’ order, we can find 14 clauses of this type coming from 10 different texts. An illustration is given in (7).

(7) þæt hy æt necstan ne magon nan land geseon.

‘that finally they cannot see any land’

(cegregdH,GDPref_1_[H]:5.20.34)

The 14 examples of the type shown in (7) correspond to 5.3% of the 266 clauses identified above and to 14.3% of all clauses with the order ‘*ne*Aux-NegYP-V’ (n = 98; XP occurs in a position other than to the immediate left of *ne*V in the remaining 85.7%). Although these frequencies are low, they are high enough to raise some initial doubts as to whether *NC-VPR* can indeed be maintained for OE.

A possible account emerges if we take a closer look at these 14 potentially problematic examples. In 7 of these, the subject is a pronoun, and in the remain-

---

*Not included here are object pronouns and subordinate clauses. The former are not likely to be diagnostics for head-final structure as they very frequently intervene between the subject and the finite verb. As for the latter, they do generally not occur between the subject and the finite verb even in clearly head-final contexts.*
ing 7, there is no overt subject in SpecTP because of subject movement (relative clauses) or because of subject omission in a second conjunct. If we focus on clauses with full DP subjects only, the result is clear-cut. Among 54 clauses with a DP subject, a negated auxiliary, a non-finite main verb, one negative YP and at least one other XP, there is not a single one that has the order ‘Su-XP-neAux-NegYP-V’. Although the amount of data allowing us to test P with full DP subjects is not huge, the fact that 54 clauses would have the potential ingredients to violate *NC-VPR but do not do so is nevertheless suggestive. HH’s conclusion is therefore supported if an analysis for the 14 clauses with a pronominal subject or no overt subject can be given in terms of a head-initial TP structure.

3.2 *NC-VPR: Apparent counterexamples

HH also identify examples with subject pronouns as the main source of potential counterarguments against *NC-VPR in OE. They suggest that subject pronouns can cliticize to C and that the element intervening between the subject and the finite auxiliary in an example like (7) either occupies SpecTP (following Pintzuk’s (1993, 1999) hypothesis that OE is a symmetric V2 language with SpecTP as a topic position) or is fronted through Stylistic Fronting. None of these assumptions is uncontroversial, however. There is no evidence suggesting that OE subject pronouns are true head clitics, the status of OE as a symmetric V2 language has generally not been accepted in the literature, and whether there are independent reasons for postulating a process of Stylistic Fronting is by no means certain.

In terms of recent approaches to the clausal syntax of OE, however, most apparent counterexamples to *NC-VPR can be accounted for quite straightforwardly. It has been widely assumed in the literature that above TP there is a second projection hosting subjects in the OE clause structure. Different labels have been given to this projection by different authors. But what is common to all these analyses is the assumption that subject pronouns always move to the higher subject position whereas full DP subjects generally occur in the lower one but can occasionally also move higher. The relevant part of the clause structure is shown in (8), which is based on the most recent version of this approach (Walkden 2017).

Interestingly, Haeberli & Haegeman (1995: 100, ex. 29b) cite such a case, but the relevant text is not included in the YCOE. Cf. fn. 5 below for a further observation concerning this example.
Walkden uses the label CP1 for the higher host of subjects, where CP1 combines the features Fin and Fam (familiar Topic) of a split CP structure. It is the Fam feature on C1 which triggers the systematic movement of subject pronouns to SpecCP1. For subordinate clauses, it can be assumed that the complementizer is merged in C1 but then raises to C2. With C1 containing a copy of the complementizer, the finite verb cannot move to C1 and remains in T. Together with the assumption that adjuncts can occur between CP1 and TP (cf. e.g. Haeberli 2000, van Kemenade & Los 2006), we get the result that ‘Su-adjunct-V’ orders or ‘adjunct-V’ orders with a no overt subject can be derived even if TP is head-initial (cf. Haeberli & Ihsane 2016: 506). An example like (7) can thus be analysed with a head-initial structure and without reference to VPR:

(9)  [CP₂ þæt [CP₁ hy þæt [ æt necstan ] [TP hy [₃T ne magon ] [VP nan land geseon ].

13 of the 14 potential counterexamples identified earlier can be analyzed along these lines as they involve AdvPs or PPs. The only example that remains problematic is the following:

(10)  þæt wæs ða ða he Iudeas nolde nan wuht læran hwæt hi don scolden
that was when he Jews not-wanted no whit advise what they do should
‘that was when he didn’t want to advise the Jews what to do’
(cocura,CP:58.443.3.3159)

Here an argument occurs between the subject pronoun and the negated auxil-
iary. Ideas cannot be argued to be a familiar topic as the referent is not mentioned before, so movement to CP1 is not an option. The only options seem to be that the adjunct position between CP1 and TP can occasionally be targeted by an argument, or alternatively that TP allows multiple specifiers. It is not entirely clear how the viability of these options could be tested and I will therefore have to leave this issue open.

4 Conclusion

HH put forward an argument in favour of the occurrence of head-initial structure in OE (or its equivalent in a system without variation in directionality) that is based on the syntax of NC and more specifically on the observation from WF that NC is not possible for a negative constituent contained within a VP that has undergone VPR to the right of an auxiliary and that therefore sequences of the type ‘Neg1-Aux-Neg2-V’ must involve head-initial structure. In this paper, I have evaluated HH’s claims on the basis of a detailed quantitative analysis and recent theoretical proposals. Among 631 subordinate clauses with at least two negative elements, a finite auxiliary and a non-finite main verb, there is only a single one (0.2%: example 10) that cannot be straightforwardly accounted for in terms of HH’s hypothesis and recent analyses of the clausal syntax of OE. Even if we define the set of relevant examples in a more restrictive way, the frequency of exceptions remains extremely low. If we look at all subordinate clauses in which a violation of *NC-VPR could have occurred (clauses with an additional XP), we get a proportion of 1 out of 266 (0.4%). Or if we focus only on clauses with Aux-Neg-V’ order and one additional XP, the rate is 1 out of 98 (1.0%). Even though the complete absence of potentially problematic cases would have been preferable, the highly exceptional status of example 10 suggests that the argument for a head-initial inflectional projection put forward by HH stands up to close quantitative and updated theoretical scrutiny.

References


Mapping the speaker’s mind

Virginia Hill

In this squib I share some thoughts about issues that need further investigation when it comes to the mapping of conversational pragmatics. In particular, based on my work with Liliane, I will point out some problems with the cartography of evaluative, evidential and epistemic modals, for which I have no solution to suggest at this time.

Haegeman & Hill (2013) and Haegeman (2014) argue for the mapping of speech acts in a separate field above ForceP (as defined in Rizzi 1997, 2004), which roughly yields the hierarchy SAP > ForceP. Vocative phrases and injunctive particles are examples of items that merge directly in the SAP area. This approach entails that the speech act field is a root clause phenomenon, since embedding occurs only at the level of ForceP and/or lower structure (i.e., assuming that the clause typing feature located in Force is the goal of the selection probe).

However, speech acts are not the only syntactic manifestations of speaker’s point of view. As argued in Cinque (1999), speech act modality is at the highest level of a hierarchy that further maps the speaker’s mind, as in (1), where other modal phrases capture evaluations, evidentiality and epistemicity, to which I will refer as the E-modal complex.

(1)  Mood_{speech act} > Mood_{eval} > Mood_{evid} > Mod_{epistemic} > TP

It is unclear how (1) can be converted to the cartographies of the left periphery proposed in Haegeman & Hill (2013), or in Rizzi (1997, 2004); see also Kidwai (2010) for a consideration of similar questions. Does the E-modal complex belong to SAP or to ForceP? The answer has consequences for understanding and delimiting the root clause phenomena.

There are already studies on the status of the E-complex, which aim to define the location of the relevant adverbs. In this sense, Haegeman (2010) points out that certain configurations, such as root and selected clauses, are more favourable to speech act adverbs and the E-modal complex, insofar as they avoid intervention effects, such as noticed with adverbial and conditional clauses.
While this is true, there is also evidence that the distribution of E-adverbs is not clear cut even when they merge in root and selected clauses. For example, the English adverb *clearly* and the Romanian counterpart *evident* may occur either above the Force complementizer, as in (2a) and (3a), or lower, as in (2b) and (3b). The evidential interpretation is the same in either position.

(2)  
   a. *Clearly that*, for whatever reason, the information wasn’t getting through on the ground. (from Radford 2013)  
   b. Tolkien’s way of thinking *clearly* spoke deeply to Lewis. (McGrath 2013: 150)

(3)  
   a. *Evident că* pe Maria cu avionul o vom trimite (nu cu clearly that DOM Maria with plane.the her-will.1PL send not with trenul).  
      train.the  
      ‘Clearly, we’ll send Maria by plane, not by train.’  
   b. Pe Maria o vom trimite *evident* cu avionul, nu cu DOM Maria her-will.1PL send clearly with plane.the not with trenul.  
      train.the  
      ‘Clearly, we’ll send Maria by plane, not by train.’

In (2a) and (3a), the adverb precedes not only the complementizer but also topic and focus constituents, which makes very unlikely an analysis that would locate the complementizer lower than Force (e.g., in Fin, as in Radford 2013 versus Force in Hill 2007). The point is that the versions in (2b), (3b) are embeddable under selection, whereas the versions in (2a), (3a) are not, as further shown in (4). Note that the evidential has a speaker oriented reading in (4a)-(4c), while in (4b)-(4d) the reading can be either speaker oriented or subject oriented. Crucially, the speaker oriented reading is not lost.

(4)  
   a. *He wrote that *clearly that*, for whatever reason, the information...  
   b. He wrote that, for whatever reason, the information *clearly* wasn’t getting through.  
   c. *Ne-a scris că *evident că* pe Maria cu avionul au to.us-has written that clearly that DOM Maria with plane.the have trimis-o.  
      sent-her
d. Ne-a scris că pe Maria au trimis-o evident cu avionul.
   ‘He wrote to us that they clearly sent Maria by plane.’

Such data are challenging for the current cartographic analyses, since it is not clear where the E-features are mapped (i.e., on SA, Force/Fin or T?).

This is not a problem only for cartographic analyses. Heavily semantic based analyses also fail to grasp the contrast in (4). For example, Krifka (2017) argues that the distinctions between the aspects involved in an assertion are not only semantic but also syntactically mapped. This is not a new idea for those working in cartography (it is, in fact, the driving principle for the development of clausal hierarchies since Rizzi 1997), but it receives novel semantic justification. Importantly, Krifka (2017) proposes a separate syntactic mapping of the semantic distinctions, as shown in (5). For a more detailed discussion of (5) see Frey (2018) and subsequent work.

(5)  ActP > CmP/JP > TP

In (5), TP is the domain of the proposition, where the truth-value is established, and which would roughly correspond to ForceP in cartography (i.e., it includes contrastive focus). Beyond TP, we deal with non-at-issueness. JP (which stands for judgment phrase) and CmP (commitment phrase) is the area where main clause operations take place (e.g., merging E-adverbs, E-related discourse particles, contrastive left dislocation). ActP (speech act phrase) also contains elements that qualify as main clause phenomena, but occur more peripherally (e.g., Hanging Topic, illocutionary particles, question tags).

In this framework, a clause is built bottom-up and may vary as to the level attained beyond TP: the presence of elements relevant to speech acts triggers the clausal projection up to ActP, while, in the absence of such elements, the derivation may stop either at CmP/JP (if material with features relevant to this domain is present in the clause), or at TP, if there is no appropriate trigger for the projection of the not-at-issue area.

Crucially, the derivation in (5) predicts that the different levels have a different distribution, a hypothesis presently explored (and confirmed) for German in Frey (2018). For example, Frey points out that a question tag can only appear with an ActP, and when this ActP is adjunct, it falls outside the structure of the
clause it relates to, as shown in (6a) versus (6b).

(6) a. *[Weil Maria sehr begabt ist, hab ich recht], wird sie schnell since Maria very talented is have I right will she quickly promovieren.

   b. Maria wird schnell promovieren, [ist sie doch sehr begabt, hab Maria will quickly graduate is she very talented have ich recht?]

   ‘Maria will quickly graduate, she is very talented, isn’t she?’

Along the same lines, a discourse particle like ja in (7b) demands that its host be at least a JP. JP has to be attached high in its host, therefore binding into an adverbial clause which contains a discourse particle is not possible.

(7) a. Weil er sehr erschrocken ist, wurde jeder bleich.

   because he very frightened was became everyone pale

   ‘Because they were frightened, everybody turned pale.’

b. *Weil er, ja sehr erschrocken ist, wurde jeder bleich.

   because he PART very frightened was became everyone pale

   ‘Maria will quickly graduate, she is very talented, isn’t she?’

Romanian brings independent confirmation for the hierarchy in (5), since the complementizer că ‘that’, obligatory with declarative complement clauses, may also occur, optionally, at the border between the fields above TP. For example, in (8), the speech act adverb ‘frankly’ and the promissive particle zău, which qualify as elements of ActP, embed a ‘that’ headed JP field containing the adverb ‘surely’, which further embeds a ‘that’ headed TP, the entire structure qualifying as a root clause (see Hill 2007 for tests verifying the mono-clausal versus bi-clausal status of such structures).

(8) Cinstit zău (că) bineînțeles (că) voi sosie la timp.

   frankly PRT that surely that will.1SG arrive in time

   ‘Frankly, I will surely arrive in time.’

So the sentence in (8) supports the field separations in (5) by showing the possibility of ‘that’ insertion in-between these fields. However, the sentence in (8) is unembeddable, as shown in (9), either as a complement or as an adjunct, as long as its level is ActP or JP, a conclusion that also follows, on independent grounds,
from cartographic analyses.

(9) a. *A has promised that frankly PRT that surely la time will.3SG arrive in timp.

b. *Deși cinstit zău că bine înțelește că va sosie la timp... although frankly PRT that surely la timp...

However, if we eliminate the ActP elements, we can get embedding on the condition that the E-adverb is somewhere lower in the structure (i.e., că is ruled out), as in (10). As mentioned for (4), embedding, as in (10a), allows for a double reading, where the E-adverb may reflect either the point of view of the speaker or of the grammatical subject (under a reportative structure), depending on the context. On the other hand, in the adjunct clause in (10b) only the speaker’s point of view is a valid option.

(10) a. A has promised that will.3SG arrive surely that in time ‘He promised to surely arrive in time.’

b. Deși va sosie bine înțelește că la timp... although will.3SG arrive surely that in time ‘Although he will surely arrive in time...’

The contrast between (9) and (10) replicates the contrast signalled in (4). Crucially, Kriřka’s (2017) proposal falls short of explaining why this would be so: Why is embedding disallowed when E-adverbs are merged high but not when they are merged low in the structure, since the mapping of the formal feature that triggers this merge must be systematically associated with the same functional head? In other words, the presence of the evidential adverbs should always signal the presence of JP, which is predicted in Kriřka’s system to systematically rule out embedding, contrary to the facts in (10).

A more promising approach seems to come from Miyagawa’s 2010 system, where C is associated not only with phi-agreement features but also with discourse (δ) agreement features. Cross-linguistic variation follows from variation in the transfer of the δ-feature set from C-to-T. Along these lines, the contrast between (2a), (3a) and (2b), (3b) would show unstable systems, where both options are in place in one single language, with δ-agr at C in (2a), (3a), and δ-agr transferred to T in (2b), (3b). This would cover the grammaticality contrast seen.
with embeddings, insofar as C with a specific set of discourse Agr is unembeddable, whereas structures where δ-Agr is transferred to T are embeddable.

Miagawa's system may seem instrumental for maintaining Cinque's (1999) hierarchy while also explaining the variation in the distribution of E-adverbs, as well as the consequences of this distribution for the syntactic behavior of the relevant structure. However, the extension of Miyagawa's analysis along these lines is problematic on other grounds: The δ-agr set at C is meant to capture the relation between topic/focus and comment/presupposition, which concerns truth value structures, not the non-at-issueness. Moreover, even if we include the E-feature sets in the δ-agr set at C, it is not clear how the cross-linguistic variation arising from C-to-T transfer may be sorted out, since the transfer may concern one set of δ-agr features (e.g., E-features) but not the other (e.g., topic/focus). This also leaves open the question of the speech act features set, which systematically blocks embedding.

I have no solution to suggest at this time for the apparently free distribution of E-adverbs as shown in (2) and (3) and their effects on clausal embedding. I only point out that this kind of data is worth investigating since it occurs quite often cross-linguistically, especially within the Romance language group. Whoever takes on this task will further Liliane’s work and deepen its significance for the field of syntax-pragmatics interface.

References

Some information-structure properties of Negative Preposing in English and Spanish
Ángel L. Jiménez-Fernández*

1 Introduction
In this squib I explore some discourse properties of a controversial operation, namely Negative Preposing. My interest on this issue is clearly and crucially marked by Liliane Haegeman’s research on it.

Negative Preposing (hereafter, NPr) is open to discussion because there are two main lines of research; for some linguists it is a subtype of Contrastive or Corrective Focus, whereas for others it is an instance of Verum/Polarity Focus Fronting (VFF). While for English there seems to be some consensus on the phenomenon, in Spanish there is a hot debate as to the specific information-structure status of this type of preposing. Based on the particular kind of focus of the fronted constituent, I hope to contribute to throwing some light on the topic. In Section 2, I discuss some properties of English NPr, taken and implemented from Haegeman (2000, 2012). In Section 3, I address the focus reading of Spanish NPr and suggest that in some discourse contexts Spanish is similar to English, whereas in other contexts it is different from an information-structure perspective.

2 English Negative Preposing
Negative Preposing has been defined as a subtype of focus fronting in the English literature (Haegeman 2012, Emonds 2004, De Clercq 2010), which involves

*This research has been partially funded by research project PGC2018-093774-B-I00 of Spain’s Ministry of Science, Innovation and Universities. I would also like to thank Liliane for her work on fronting and root phenomena, which have always guided my own research.
movement of a negative or non-assertive constituent to the left periphery of the sentence, by means of which the polarity of the sentence is affected yielding a negative sentence, as in (1):

(1) a. [On no account] could she move to Paris. \(\text{(De Clercq 2010)}\)
b. [Not a bite] did he eat. \(\text{(Green 1976: 384)}\)

Both adjunct PPs and argument DPs can be fronted in this type of constructions. Among the defining properties of English NPr are the subject-auxiliary inversion attested in (1) and the emphasis on the negative polarity of the sentence. The first trait has been claimed to be common to all types of focus fronting, which has led linguists to argue in favour of an analysis of NPr as triggered by a focus feature. However, clear cases of focus fronting in English do not require subject-auxiliary inversion:

(2) THIS BOOK I don’t need (but that one I do). \(\text{(Haegeman 2012: 8)}\)

Note that if no subject-auxiliary inversion takes place in NPr, the result is fully ungrammatical (though see Haegeman 2000 for a topic reading of negative elements when they exhibit no inversion):

(3) a. *[On no account] she could move to Paris.
b. *[Not a bite] he ate.

The second property that describes the phenomenon under study is the emphasis on the negative polarity of the relevant sentence caused by the fronting of the negative constituent. Sentence negation is involved in NPr. Haegeman (2000) and De Clercq (2010) argue that if a negative tag can be added to a sentence it is because this particular sentence is negative. If this is correct and sentences with NPr are negative, cases of NPr are expected to accept a neither-tag (since neither-tags are only compatible with negative clauses). This prediction is borne out in light of the data in (4):

(4) On no account could she move to Paris, and neither could Jane. (adapted from De Clercq 2010)

Haegeman (2012) provides an analysis of NPr based on the movement of the negative constituent to a designated Focus Phrase in the left periphery. This movement is triggered by a focus feature. Since what is emphasized is the negative polarity, I assume for English that this focus feature is connected with a neg-
ative feature. In other words, the triggering featural array is \[+\text{focus, } +\text{negation}\], in the spirit of decomposition in features of discourse categories proposed in Jiménez-Fernández (2015).

Support for the claim that English NPr involves focus comes from the fact that it can be used as an answer to a question satisfying its information request. This is clearly stated in Culicover (1991) and Haegeman (2000), and illustrated in (5):

(5) A: Did you see anyone?  
    B: No, NOT A SINGLE PERSON did I see.  \(\text{\textcopyright Culicover, 1991}\)

Note, however, that this is a polar question and the answer fulfills the information request on the polarity, so it is interpreted as negative. A couple of examples involving a \(\text{wh}\)-question follow, which can perfectly be replied by a NPr construction (Culicover, 1991):

(6) A: When did you ever see such a thing?  
    B: NEVER have I seen such a thing.

(7) A: When would you ever have agreed to visit Robin?  
    B: AT NO TIME would I ever have agreed to visit Robin.

Importantly, what is clear is that the fronted negative constituent has some sort of focus interpretation, based on the licit marking of it as Information Focus (IF) in (6)–(7).

Alongside the IF reading, the relevant literature associates NPr with Contrastive Focus (CF) in English (Haegeman, 2012). I have shown evidence that NPr can be interpreted as IF. Next question is whether evidence can be found supporting a possible interpretation of NPr as CF. Since Jespersen (1937), it is widely acknowledged that CF can be developed by means of clefting (E. Kiss, 1999, Frascarelli, 2000, Belletti, 2005, Haegeman et al., 2014, Cruschina, 2015). In some languages there is the option of developing CF either via fronting or via clefting, as illustrated for Italian, English and German (examples taken from Cruschina, 2015):

(8) [Context: I have met Charles]  
    a. GIANNI ho incontrato / È GIANNI che ho incontrato. \(\text{\textcopyright Italian}\)  
    b. JOHN I met / It was JOHN that I met. \(\text{\textcopyright English}\)  
    c. HANS habe ich getroffen / Es war HANS, den ich getroffen. \(\text{\textcopyright German}\)
If NPr is a subtype of CF, the prediction is that the fronted negative element should be possibly used in a cleft. Examples in (9) confirm the contrastive flavour of negative constituents, either adjuncts (9a) or arguments (9b)-(9c).

(9)  

a. It is AT NO TIME that I would ever have agreed to visit Robin.

b. It is NO INTEREST that they show in syntax.

c. It is NO ONE that I saw.

From this empirical reasoning it can be safely concluded that English NPr has two possible interpretations, namely either that of purely information focus or that of contrastive focus. Both interpretations favour an analysis in which the fronted element undergoes movement to the specifier of the designated category FocP in the left periphery.

3 Spanish Negative Preposing

Next I turn to Spanish. NPr in Spanish has not been studied in depth. Bosque (1980) notices a type of fronting which yields a negative interpretation of polarity in Spanish. The author calls these fronting operations anteposiciones negativas 'negative preposing', which I illustrate in (10a) for adjuncts and (10b) for arguments, from Bosque (1980: 34–35):

(10)  

a. En modo alguno se puede tolerar tal actitud.

b. De ninguno de esos problemas trató la reunión.

These are cases of adjunct or argument NPr, which clearly induce a negative interpretation of the sentence polarity, given the corresponding sentences in (11) with no fronting but with the explicit occurrence of the negative adverb no 'not':

(11)  

a. No se puede tolerar tal actitud en modo alguno.

b. La reunión no trató de ninguno de esos problemas.
This trait supports the idea that sentences involving NPr are marked as negative in Spanish. In the absence of fronting, some other mechanism such as insertion of the negative adverb is obligatory. The rule is known as Neg-shift (for discussion on this rule in different views, see De Clercq 2010, Haegeman 2000, Bosque 1980, Tubau 2008, Zeijlstra 2004). The test of neither-tags provided by Haegeman (2000) for English is easily extended to Spanish, thereby showing that negation in NPr applies to the whole sentence:

(12) En modo alguno se puede tolerar tal actitud, ni tampoco tu respuesta.

‘By no means can such an attitude be tolerated, and neither your answer can.’

The fact that a but-tag can be added to a NPr construction also suggests that the sentence is negative, given that but-tags are only compatible with a previous negative polarity sentence (Etxepare & Uribe-Etxebarria 2008); it may have a bound focus reading, suggesting that this is a subtype of CF:

(13) A NINGÚN ESTUDIANTE me encontré por la facultad, sino a profesores.

‘I found no students at the faculty, but I found professors.’

Concerning the formal analysis of NPr in Spanish, there are two main lines of research. On the one hand, Gallego (2007) and Battlori & Hernanz (2014) argue that this type of fronting are cases of mild focalization or weak focus fronting, suggesting that there is a focus feature triggering movement of the negative constituent and affecting the sentence polarity. Properties such as obligatory subject-auxiliary inversion are taken as evidence for their analysis.

On the other hand, Leonetti & Escandell-Vidal (2009, 2010) and Escandell-Vidal & Leonetti (2014) claim that NPr is a subtype of Polarity or Verum Focus Fronting which makes polarity negative (illustrated in (14), alongside cases of Quantifier Fronting (Quer 2002), Resumptive Preposing (Cinque 1990), etc. (see Jiménez-Fernández 2015 for a classification of types of foci in Spanish)).
Nada tengo que añadir a lo que ya dije en su día.

'I have nothing to add to what I said at the time.'

For Escandell-Vidal & Leonetti there is no information-structure partition in these constructions and no contrastive focus of the fronted element is involved. Rather, it is the sentence polarity that is emphasized. In particular, NPr highlights the sentence negative polarity, whereas Quantifier Fronting and Resumptive Fronting makes affirmative polarity more prominent (see also Hernanz 2006). This is why I am taking the three phenomena as different.

To throw some light on the possible solution for the type of focus which Spanish NPr instantiates, I will test whether it may have an Information Focus and/or a Contrastive Focus interpretation. Starting with the IF reading, Escandell-Vidal & Leonetti (2014) explicitly address the impossibility of using NPr as reply to information-seeking questions, providing an example with QP-fronting:

A: ¿ Crees que tengo interés?
   ‘Do you think I have interest?’

B: ¡Mucho interés tienes tú...
   ‘A great interest you have...’

Extrapolating the IF interpretation to NPr, note the following natural dialogue, suggesting that QP-fronting and NPr are distinct phenomena, sharing some properties such as the emphasis on the sentence polarity, but also showing distinct features:

A: ¿Qué tienes que añadir a lo que expusiste ayer?
   ‘What do you have to add to what you presented yesterday?’

B: NADA tengo que añadir.
   ‘I have to add nothing.’
This confirms that NPr may have an IF reading in Spanish as well, similar to English.

As regards the CF interpretation, let’s check whether NPr alternates with clefting. Clefting in Spanish has been held to express either IF or CF. Following Moreno Cabrera (1999) and, more recently, Feldhausen & del Mar Vanrell (2015), at least one of the interpretations of a cleft sentence is that of CF, as illustrated in (17), from Feldhausen & del Mar Vanrell (2015), adapted from Moreno Cabrera (1999: 4251):

\[(17) \quad \text{Es be-PRES-3SG } \text{Juan the that come-PRES-3SG (and not Sebas)}\]

‘It is Juan that will come (and not Sebas).’

The question is whether clefting or pseudo-clefting is possible with the constituents which are fronted via NPr. It must be noted first that clefts behave like monoclusal sentences as far as negation is concerned and the negative element requires the adverb no ‘not’ to be inserted in initial position (Bosque 1980). In doing so, the results of clefting in Spanish are fully grammatical, confirming the possible use of NPr with a corrective flavour:

\[(18) \quad \begin{align*}
A: \quad & \text{Seguro que te encuentras con algún vecino, ¿verdad?} \\
& \text{Sure that you find-PAST.2SG with some neighbour true} \\
& \text{‘For sure you met some neighbour, right?’}
\end{align*} \]

\[(19) \quad \begin{align*}
B: \quad & \text{No fue A NINGÚN VECINO a quien me encontré} \\
& \text{not be-PAST.3SG to no } \text{neighbour whom me find-PAST.1SG} \\
& \text{en la entrada, sino a tu novia.} \\
& \text{in the hall, but to your girlfriend.} \\
& \text{‘It wasn’t any neighbour at all who I met in the hall, but your girlfriend.’}
\end{align*} \]

\[(19) \quad \begin{align*}
& \text{No fue DE NINGUNO DE ESOS PROBLEMAS de lo que} \\
& \text{not be-PAST.3SG of none of those problems of the that} \\
& \text{trató la reunión.} \\
& \text{treat-PAST.3SG the meeting} \\
& \text{‘It wasn’t any of those problems that the meeting discussed.’}
\end{align*} \]

Just in passing, Herrero (1992: 207) states that in contexts of special expressivity such as replies or rejections negative elements can be focalized in a (pseudo-)cleft. In this connection, correction and contrast also represent especially expressive discourse situations; note that the cleft alternates with NPr in (B-B’):
From the discussion on Spanish NPr it can be drawn the conclusion that this phenomenon can be interpreted as either IF or CF, exactly as in English. This does not mean an exhaustive list of interpretations. To the contrary, Spanish NPr may also have a VFF reading. To accommodate the data, Spanish NPr involves movement of a negative element to the left periphery.

Escandell-Vidal & Leonetti (2014) mention the possibility that this movement is not motivated by any triggering feature, but they do not elaborate on this. I assume that there is a feature causing the fronting operation. This feature is similar to the one proposed for English, but differs in that the [+focus] feature is diminished to just [+emphasis] in those cases where the NPr phenomenon is interpreted as VFF. Thus Spanish NPr displays a [+emphasis, +negation], a combination in charge of emphasizing the negative polarity of the whole sentence. The precise syntactic position targeted by the negative element is a pending issue in my current research on the topic (see Jiménez-Fernández 2018).

References


Negative adjectives

Wim Klooster

1 Introduction

Liliane's work, especially her study of the syntax of negation (Haegeman 1995), has been very inspirational to me. Also, her *Introduction to Government and Binding Theory* was for many years a useful and instructive guide for my students (and me!), thanks to its clarity of style and exposition.

I chose 'negative adjectives' as subject of this squib – based on Dutch and English examples – because I believe it is one of the few not treated in her work on negation.

The term 'negative adjectives' may refer to either adjectives functioning as Negative Elements (NEs), or to antonyms of positive adjectives. (There are adjectives without antonyms, but I will leave that point aside here.) In this contribution, I will attempt to challenge the idea that if an adjective acts as an NE, it can't be a positive antonym. I will concentrate on gradable adjectives.

Examples with adjectival NEs are given below (with Negative Polarity Items (NPIs) under-lined):

(1) Het is *gevaarlijk* er ook maar iets over los te laten.
    'It is *dangerous* to let on anything at all about it.'

(2) Hij vond het *moeilijk* ook maar iets toe te geven.
    'He found it *hard* to admit anything at all.

An adjective's having a negative prefix (*impolite, unpleasant*) does not mean that it is potentially an NE, nor that it is a negative antonym (henceforth, a '[-Pol] adjective'). As illustrated in (1), it is *gevaarlijk* 'dangerous', not *ongevaarlijk* 'safe', lit. 'undangerous', that licenses NPIs. And, as will be argued below, *dangerous* and *hard* are positive (henceforth, ' [+Pol]') adjectives.

Several tests have been suggested in the literature for determining which of a pair of antonymous adjectives is [+Pol]:

146
First, if a member of a pair of antonyms, e.g. *lang/kort* ‘long, tall’/’short’, takes measure phrases, it is positive: *1 meter lang* vs. *1 meter kort*.

Second, comparative constructions with inchoative copulas, like X gets/becomes [Adj] [+Pol]-er, have paraphrases with (syonyms of) *stijgen* ‘rise’ and *dalen* ‘fall’, corresponding to [+Pol] and [−Pol], respectively. The adjectives in question (say, *warm/cold, or expensive/cheap*) should be linkable to objective parameter nouns, i.e. nouns referring to measurable properties (e.g. *temperature or price*):

(3) X wordt warmer/kouder – De temperatuur van X stijgt (neemt toe)/daalt (neemt af)
    X gets warmer/colder – X’s temperature rises (increases)/falls (decreases)

(4) X wordt duurder/goedkoper – De prijs van X stijgt/daalt
    X is getting more expensive/cheaper – The price of X is rising/falling

This test does not work for adjectives not related to objective parameter nouns. For instance, the pair *schadelijk/onschadelijk* (harmful/harmless) lacks a corresponding objective parameter noun suitable for an unambiguous test. The ‘degree of harmfulness’ just as well as the ‘degree of harmlessness’ can be said to increase or decrease.

Third, whereas speakers have little difficulty in interpreting expressions like ‘zero fertility’ or ‘zero speed’, they find it hard to make sense of ‘zero infertility’, or ‘zero slowness’. While many positive gradable adjectives can be associated with some ‘zero point’ (absence of any degree of e.g. fertility or speed), it appears difficult to determine such a point for their negative antonyms. (But see the proviso in the last section, in connection with the ‘black-and-white’ effect.) In Sassoon’s (2010) terms, while negative adjectives map entities to values that are ‘linearly reversed’ and ‘linearly transformed’ in comparison with their values in their positive antonyms, we do not know which linearly reversed function they denote. Their ‘zero point’ is undetermined.

That an adjective licensing NPIs is not necessarily [−Pol] can be demonstrated with, for instance, *moeilijk* (hard, difficult). As we saw above, it may occur as a (weak) NE. Still, it must be considered the positive antonym of *makkelijk* (easy): *zero easiness* is puzzling at best, whereas *zero difficulty* is readily interpretable. Not surprisingly, of the two parameter nouns *moeilijkheidsgraad* (degree/level/rate of difficulty) and (ge)makkelijkheidsgraad (ditto of easiness), the former is more natural and common. In addition, a *low degree of difficulty* is applicable in
cases where a thing can be qualified as easy – which is an indication of the ‘neutral’ character of degree of difficulty, even though there is no standard measure of difficulty. A low degree of easiness, on the other hand, apart from sounding odd, does not appear to cover things that may be called difficult.

In order to get a clearer view of what is involved in determining the polarity of gradables, it may be useful to consider their classification into subsets. This will be done in the next section.

2 A taxonomy of gradables

The following is a summary, with some slight adaptations, of a classification into subsets of gradables proposed in Klooster (1976).

Gradables can be either [+Pol] or [−Pol]. They are, furthermore, either ‘objective’, in the sense that they can be linked to objective (i.e measurable) parameters like length or luminosity, – or ‘subjective’. A subset of the objective gradables consists of adjectives bearing the feature [+Oriented], while all others (whether subjective or objective) are [−Oriented]. (For the sake of convenience, I adopt Bierwisch’s terminology here.) [+Oriented, +Pol] adjectives allow equative constructions with half as, twice as etc., whereas their antonyms do not; [−Oriented] adjectives on the other hand, whether [+Pol] or [−Pol], always allow such constructions (Bierwisch 1967):

[+Oriented]:

(5) a. The table is twice/half as long as the bench.
   b. *The table is twice/half as short as the bench.

(6) a. He is twice/half as old as his brother.
   b. *He is twice/half as young as his brother.

[−Oriented]:

(7) a. John is twice/half as good at tennis as Bill.
   b. John is twice/half as bad at tennis as Bill.

(8) a. The room is twice/half as light as the corridor.
   b. The room is twice/half as dark as the corridor.

The [+Pol, +Oriented] adjectives can all be used in a neutral sense, that is, without presupposing some subjective norm. A subset of these take measure phrases
To determine whether a given adjective is subjective, one can check if it fits in the context I find X ___-er (more ___ ) than Y. here are subjective adjectives for which this test is not suitable, such as dizzy or sick. However, these can be tested in contexts like I feel ___-er (more ___ ) now than I did before. I will refer to these tests as ‘subjective comparative tests’.

Objective gradables can be used in a subjective sense, presupposing some subjective norm. Examples are John is old, You are still young, It’s warm/cold in here. They do not, however, pass the first subjective comparative test mentioned above, and are always associated with some objective scale of measurement. Old, though allowing use in the subjective sense, may occur in the ‘neutral’, objective sense as a measure adjective or, for instance, in questions like How old are these kittens? Young, presupposing a norm in the example above, may also occur in the ‘objective’ sense: a few months too young. Similarly, subjective warm (warmₙ) has an objective counterpart (warm₀). Like old, warm₀ can be related to an objective scale of measurement. Its antonym cold, too, may occur in either sense: it’s cold outside vs. 20 degrees too cold for April.

The above is summarised in (9):

<table>
<thead>
<tr>
<th>SUBJECTIVE</th>
<th>OBJECTIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>+POL</td>
<td>light, warmₙ, etc.</td>
</tr>
<tr>
<td>−POL</td>
<td>dark, cold₀, etc.</td>
</tr>
</tbody>
</table>

(Measure adjectives are in bold italics)

## 3 Subjectivity, negative antonyms

The meaning aspect shared by all objective gradables can be described by means of a scale of indeterminate length, starting at a point 0, with points on the scale

---

1 They do pass the second one, but only when taken in a special sense, as in I feel (c)older now than I did before.

149
representing certain values. In (10), the value \( p \) represents the height of, say, some poplar tree, and \( e \) the height of some elm tree, where the poplar is taller than the elm.

\[
\begin{array}{c}
0 \quad e \quad p
\end{array}
\quad \text{(parameter of height)}
\]

The scale being directional, we say that \( p \) is at positive distance from \( e \), or \( D(p, e) \). For expressing the reverse, negative distance, we need the complement of \( D \), \( D' \); for all parameter values \( x, y \) \((x \neq y)\), \( D'(x, y) \) is equivalent to \( D(y, x) \) as well as to \( \text{NOT}(D(x, y)) \). Thus, since \( e \) is at a negative distance from \( p \) (the elm is less tall than the poplar), we can also represent the relation between \( e \) and \( p \) by writing \( D'(e, p) \).

One difference between objective and subjective gradables seems to be that the scales associated with the latter do not have a definite starting point or ‘zero’ point. The fact that the ‘zero point’ criterion mentioned in section 1 gives the clearest results with objective adjectives (e.g., ??zero slowness vs. ??zero carelessness) may well be related to this. According to Sassoon (2010: 176), “positive [objective, WK] adjectives tend not to have a maximal point (e.g., there is no tallest point), a fact which renders the zero point of their negative antonym undefined.”

For subjective gradables, as well as for [–Pol] objective ones like short or dark, we must introduce a norm \( N \) on the scale, the scale having, in the case of objective adjectives, a zero point, as in (10). \( N \) is not necessarily a point; it may be a ‘grey area’ corresponding to notions like ‘average’ or ‘neither’ (e.g., ‘neither beautiful nor ugly, but something in-between’). Thus, we can represent the relevant semantic properties of, say, beautiful/ugly as in (11), where \( x, y, u, v, w \) and \( z \) are ‘beauty values’ attributed to \( X, Y, U, V, W \) and \( Z \), respectively.

---

2 The well-known fact that than-clauses contain an abstract negative operator can be accounted for in this approach by having the logical representation of comparative constructions express that there are values \( v \) between the compared values, say, \( e \) and \( p \), such that \( D(v, e) \& \text{NOT} D(v, p) \) (Klooster 2016). In terms of ‘extent’ this can be reformulated by stating that, for some extents \( e^p, e^e \), the poplar is tall to \( e^p \) while it is not the case that the extent \( e^e \) to which the elm is tall equals or includes \( e^e \).
In (11), both \( x \) and \( y \) are at negative distance from \( N \) (‘\( X \) and \( Y \) are ugly’). Furthermore, \( x \) is at negative distance from \( y \) (‘\( X \) is uglier than \( Y \)’), and \( y \) is at positive distance from \( x \) (‘\( Y \) is less ugly than \( X \)’). The relationships between \( w \), \( z \) and \( N \) can be stated analogously, with ‘negative’ switched to ‘positive’ and vice versa.

If two values in a comparison are within \( N \), like \( u \) and \( v \) in (13), then, in my judgment, something like ‘(both are average but) \( U \) is less beautiful than \( V \)’ would seem more appropriate than ‘(both are average but) \( V \) is less ugly than \( U \)’. If this intuition is correct, it is consistent with the idea that we are dealing with a ‘scale of beauty’ rather than of ugliness, and hence, with the intuition that \( \text{beautiful} \) is \([+\text{Pol}]\).

What if the compared values are each in a different ‘subparameter’, \( \text{ugly} \) and \( \text{beautiful} \)? In a context where, say, Snow White is beautiful and the queen (in disguise) ugly, \textit{The queen is less beautiful than Snow White} would perhaps be acceptable as an understatement. Substituting \textit{less beautiful} by \textit{uglier} here, however, would produce a statement not applicable in such a context. I am not quite certain about \textit{Snow White is less ugly than the queen} and \textit{Snow White is more beautiful than the queen}. The former seems more comical than the latter.

Can there be a ‘zero’ beauty value? On the Web, occurrences – in the senses intended – are scarce of \( \text{zero beauty} \) (three times), as well as of \( \text{zero ugliness} \) (just once). It seems safe to assume that subjective adjectives do not have minimum or maximum values. It could very well be, though, that whenever \( \text{zero beauty} \), \( \text{zero difficulty} \) etc. do occur, such expressions refer to the left hand boundary of \( N \), or to \( N \) itself, should \( N \) be a single point on the scale. The latter occurs in certain contexts where \( N \) reduces to a sharp boundary between the two subparameters, giving rise to the ‘black-and-white’ effect, so that \textit{not beautiful} comes to mean ‘ugly’ and not \textit{ugly}, ‘beautiful’, etc. In such cases, expressions like zero ugliness do make sense, contrary to what the third criterion mentioned in section \[1\] would lead one to expect.

In any case, clearly the notion ‘norm’ must be present in some form or other in lexical entries of subjective gradables and \([-\text{Pol}]\) objective ones. The entries of all \([-\text{Pol}]\) adjectives, furthermore, should somehow express that they map
some value $x$ on the scale involved, such that, for some norm $N$, $x$ is at a negative distance from $N$, that is, \( \not\in D(x, N) \).

The fact that pairs like \textit{warm/cold}, \textit{light/dark}, or \textit{sharp/vague} (said of images) refer to sensory perceptions, which are in themselves subjective, may be related to their sharing the feature \([-\text{Oriented}]\) with subjective adjectives. At the same time, like other objective adjectives, they are associated with objective parameters (temperature, luminosity, resolution).

4 Conclusion. Negative antonyms vs. adjectival NEs

Kennedy (1998) argues that there is a connection between a gradable adjective’s being \([-\text{Pol}]\) and its being monotone decreasing, that is, an NE. But as I argued above, this connection does not seem to exist. \textit{Dangerous} in \(\text{(1)}\) is \([+\text{Pol}]\): ‘zero danger’ seems less odd than ‘zero safety’. (Occurrences of zero safety can be found on the Web, but never with the meaning ‘absence of safety’.) \textit{Difficult} in \(\text{(2)}\), as pointed out earlier, is also \([+\text{Pol}]\). The fact that nevertheless these adjectives license NPIs must mean that it is not the negative distance from some norm $N$ which somehow causes them to function as NEs. Rather, the evaluation of properties like ‘dangerous’ or ‘difficult’ relative to what is desirable underlies their being felt as negative.

Of the antonyms \textit{zwaar} ‘heavy, hard’ and \textit{licht} ‘light, easy’, the former is \([+\text{Pol}]\). Yet it can function as an NE, while \textit{licht} cannot (cf. \(\text{(12)}\)). Similarly, \textit{ver} ‘far’ is \([+\text{Pol}]\) – for instance, it takes measure phrases – and \textit{dichtbij} ‘close’ is \([-\text{Pol}]\). But \textit{far} licenses NPIs, whereas \textit{close} does not (cf. \(\text{(13)}\)).

\(\text{(12)}\)

\begin{enumerate}
\item a. Het viel hem \textit{zwaar} om dat te verkroppen.
  \quad It fell him heavy \textbf{COMP} that to swallow
  \quad ‘It was \textit{hard} for him to swallow that.’
  
  b. *Het viel hem \textit{licht} om dat te verkroppen.
\end{enumerate}

\(\text{(13)}\)

\begin{enumerate}
\item a. De \textit{hut} was \textit{ver} van (\textit{ook maar}) \textit{enige} \textit{bewoonde plek}.
  \quad The \textit{hut} was \textit{far} from even \textbf{just} \textit{any} inhabited \textbf{COMP} spot
  \quad ‘The \textit{hut} was far from \textit{any} inhabited place.’
  
  b. *De \textit{hut} was \textit{dichtbij} van (\textit{ook maar}) \textit{enige} \textit{bewoonde plek}.
\end{enumerate}

The ‘negativity’ of \textit{zwaar} ‘hard’ and \textit{ver} ‘far’ in \(\text{(12)}\) and \(\text{(13)}\) may have an explanation similar to the one suggested above regarding \textit{dangerous} and \textit{difficult}.

Concluding, I submit that the above observations in connection with criteria
for determining whether an adjective is ‘negative’ (in the sense of [–Pol]) and
those regarding adjectival NEs, at least cast doubt on the idea that there are
adjectives licensing NPIs by virtue of their being negative antonyms.

References

Bierwisch, M. 1967. Some semantic universals of German adjectives. *Founda-
tions of Language* 3. 1–36.

Press.

semantics.uchicago.edu/kennedy/docs/monotonicity.pdf](http://semantics.uchicago.edu/kennedy/docs/monotonicity.pdf).

& A. Evers (eds.), *Lijnen van taaltheoretisch onderzoek*. 229–259. Groningen:
Tjeenk Willink.

Klooster, W. 2016. NEG CPs and monotonicity. In A. Bannink & W. Honselaar
(eds.), *From variation to iconicity. festschrift for Olga Fischer on the occasion of

Sassoon, G.W. 2010. Degree functions of negative adjectives. *Natural Language
Semantics* 18. 41–81.
Ik richt maar even het woord tot jou in onze moedertaal, hoewel... dat klopt natuurlijk niet helemaal: allebei zijn we opgegroeid met de ronkende klanken van ons heerlijke West-Vlaamse dialect, op de cadans van het alom tegenwoordige motto – de provinciale hymne zou je bijna kunnen zeggen – doe mo geweune voart ‘doe maar gewoon voort’.

En ‘gewoon voortdoen’, dat heb je gedaan. . . en hoe! Je bent een van de allerbeste internationale taalkundigen, met een duizelingwekkende wetenschappelijke output op zoveel domeinen van de syntaxis en over zoveel talen. Het is dan ook niet makkelijk om alles wat je gedaan hebt en alles waar we je dankbaar voor zijn in één tekst te vatten. Toch zal ik hier een poging wagen, omdat ik je graag mee wil geven hoezeer ik jou en je taalkundige werk apperceiveer, en hoezeer je mij en vele anderen inspireert.

Wat mij steeds opnieuw treft wanneer ik een artikel of boek van je lees, is hoe fundamenteel eerlijk je onderzoek is. Die eerlijkheid manifesteert zich eerst en vooral in de stevige empirische basis van je analyses: generalisaties en claims baseer je op reeksen van descriptieve argumenten, mooie minimale paren en sprekende voorbeelden van authentiek taalgebruik uit verschillende registers. In jouw artikels worden lastige of onverklaarbare gegevens niet in voetnoten wegestopt; in jouw artikels zijn er geen theorie-interne argumenten zonder onafhankelijke descriptieve staving. Deze empirische onderbouwing maakt je werk tijdloos, zorgt ervoor dat het relevant is – en zal blijven – voor huidige en toekomstige taalkundigen van verschillende strekkingen.

* [English translation below]
Je absolute en principiële eerlijkheid is daarnaast ook duidelijk voelbaar in de manier waarop je krediet geeft aan andere onderzoekers. De bibliografieën van je artikels en boeken zijn nagenoeg exhaustief, je verwijst steevast naar een groot aantal eerdere werken over het onderwerp, of dat nu recente of oudere publicaties zijn, in grote of minder grote tijdschriften, van mensen uit je inner circle of van onderzoekers met een volledig andere achtergrond, van analyses waar je het mee eens of niet mee eens bent. Je neemt iedereen serieus en behandelt iedereen gelijk, of het nu gaat om een absolute beginner of een bekende naam in het wereldje.

Vanuit een meer theoretisch perspectief hou ik ook erg van de manier waarop je, met een hele rij aan glasheldere descriptieve en theoretische argumenten, nagenoeg vastgeroeste hypotheses in de taalkundige literatuur aan het wankelen weet te brengen, of het nu gaat om de positie van links-perifere topics in het Engels en de Romaanse talen, of de syntactische positie van de focus van een gekloofde zin, om er maar een paar te noemen. Als jij je licht hebt laten schijnen op een of ander fenomeen, wordt er nadien niet meer geschreven ‘there is no consensus on...’ de argumenten die je aanbrengt zijn meestal zó overtuigend en doordacht dat er nadien gewoonweg geen discussie over het onderwerp meer is.

Je onderzoek heeft veel andere taalkundigen geïnspireerd: je werk zet steeds aan tot nadenken, het geeft de lezer zin om iets verder uit te spitten, te toetsen aan een andere taal of dialect. Het is een grote verdienste van je onderzoek, dat het tegelijk iets afondt en opnieuw een opening creëert en toekomstig onderzoek initieert. Zo ben ik een hevige fan van je werk over gekloofde zinnen, waar je onomstotelijk aantoont dat het gekloofde element zich niet in de linker periferie van de zin bevindt. Dat is alvast één debat minder in de taalkundige literatuur. Ook je vele onderzoek naar bijwoordelijke bijzinnen en hoofdzinsverschijnselen, waarin je de functioneel-taalkundige literatuur de erkenning geeft die ze verdient in dit domein, en waarin je het verband legt met lokaliteitsrestricties, heeft mijn werk over het Frans erg beïnvloed. Verder hou ik van de manier waarop je concepten uit Informatiestructuur gebruikt in je syntactisch werk: met een grote omzichtigheid, en zonder ze volledig te herleiden tot een puur syntactisch kenmerk. Met je ideeën over voorwaardelijke bijzinnen heb ik tot nog toe niets gedaan, maar dat zit er zeker aan te komen. Je blijft inspireren.

Zelf blijf je ook inspiratie vinden: je hebt een immense gedrevenheid, een bijna buitenaardse of in ieder geval toch bovenmenselijke output (vergeef mij deze adjectieven waarbij je – nuchter als je bent – wellicht je wenkbrauwen fronst). In Leuven ging en gaat nog steeds het verhaal – ik vertelde het je al eens eerder –
dat er ooit een scatterplot werd gemaakt van de wetenschappelijke output van Vlaamse taalkundigen, en dat jij daarin een heel extreme waarde innam. In die zin hebben de adjectieven buitenaards en bovenmenselijk, die ik hierboven gebruikte, toch wel enige iconische waarde.

Het eerste dat ik van jou las, was je *Introduction to Government and Binding Theory*, het boek met de grijze kaft dat ik na een tijdje bijna uit het hoofd kende, een handboek van een ongekende helderheid, enorm aantrekkelijk voor de beginnende en hongerige taalkundige die ik was. Met dat boek heb je een hele generatie jonge mensen overtuigd van het belang van de generatieve taalkunde, heb je glashelder uitgelegd wat de empirische basisvragen zijn van het model, de explicative power ervan, de basismethodologie, de uitgangspunten. In naam van al deze mensen: bedankt daarvoor!

Fantastisch is het hoe je de generatieve taalkunde weer helemaal binnenbracht in België. Ik weet nog goed hoe we, zo'n tien jaar geleden, hoorden van je benoeming als Odysseus-professor in Gent. Het was alsof er plots weer een toekomst was voor generatieve taalkunde in Vlaanderen: wij onderzoekers die ons tot dan toe een beetje ‘verscholen’, of onderdak zochten in het buitenland, mochten plots weer gewoon generatieve-taalkundige-in-Vlaanderen zijn . . . zo voelde het toch aan. Met jouw aanstelling was er in Vlaanderen plots weer openlijk erkenning van de generatieve taalkunde; plots was er weer bedrijvigheid, waren er conferenties, contacten en mogelijkheden, en werden er weer heel wat jonge mensen aangetrokken om hier taalkundig onderzoek te doen over fascinerende syntactische onderwerpen. Bedankt ook daarvoor!

Je hebt ook het West-Vlaams op de taalkundige kaart gezet: je artikels over complementizer agreement, negative doubling, possessor relations en pronomina in het West-Vlaams zijn pareltjes die – voor wie er nog aan zou twijfelen – aantonen hoe schijnbaar kleine observaties op een micro-niveau implicaties kunnen hebben voor taalvergelijking op een macro-niveau en taalkundige theorie in het algemeen. Dankzij jou is ‘West-Flemish’ een begrip geworden in de generatieve taalkunde: hoe vaak heb ik niet op een of ander diner van een conferentie, waar iemand me vroeg naar mijn moedertaal, zonder veel verdere uitleg kunnen antwoorden: “Liliane’s dialect”.

Bedankt, Liliane, dat je met zoveel zorg zetelt in wetenschappelijke commissies. Voor veel mensen, ook niet-generativisten, is het een geruststelling dat iemand met zo’n principes, iemand die zo objectief en eerlijk is, waakt over de ingediende dossiers.

Bedankt, voor de kansen die je me hebt gegeven, bijvoorbeeld om te spreken op de *Subjects Workshop* in 2012. Bedankt ook voor de kleine babbeltjes die we
zo nu en dan hadden, meestal in de marge van een lezing, doctoraatsverdediging of congres, soms via de telefoon, en waarbij het me meer dan eens opviel hoe je terecht verontwaardigd blijft over bepaalde situaties, zonder na een tijdje over te gaan tot gelaten schouderophalen.

Het is erg moeilijk om voor te stellen dat jij op emeritaat gaat, en stiekem denk (en hoop) ik dan ook dat je gewoon even productief zult blijven als voor-dien, omdat je je nu helemaal kunt wijden aan de kern, het taalkundig wetenschappelijk onderzoek, zonder alle administratieve rompslomp die een academische carrière ook met zich meebrengt. Als dat is wat je wil, dan wens ik het je van harte toe!

Maar toch hoop ik dat je daarnaast ook substantieel meer tijd zult krijgen én nemen voor al het niet-taalkundige dat je graag doet, voor alles wat je rust en warmte brengt, voor de kleine dingetjes die het leven zo mooi maken, van frozen yoghurt en poezengesnor tot intense momenten met wie je liefhebt. Bedankt voor alles! Toet in’ droai!

Karen Lahousse (KU Leuven, België)
Dear Liliane,

I’ll address you in our mother tongue, although... that’s not exactly true, is it? Both of us grew up with the throbbing sounds of our delightful West-Flemish dialect, on the cadence of the omnipresent motto – the provincial hymn, one could say – *doe mo geweune voart* ‘just keep going’.

And ‘just keep going’, you’ve certainly done that... and how! You are one of the very best international linguists, with a dazzling scientific output in so many areas of syntax and about so many languages. This makes it quite hard to summarize in a small text like this one everything you’ve done and all that we’re so grateful for. I will nonetheless give it a try, because I would really like to show you how much I appreciate you and your linguistic work, and how much you inspire me and others.

What strikes me time and again when I read one of your articles or books is how fundamentally honest your research is. That honesty manifests itself first and foremost in the thorough empirical basis of your analyses: you base your generalizations and claims on descriptive arguments, elegant minimal pairs and vivid examples of authentic language use from different registers. In your articles, tricky or inexplicable data are not hidden in footnotes; your articles don’t contain theory-internal arguments without independent descriptive support. This empirical foundation makes your work timeless and assures that it is – and will remain – relevant for contemporary and future linguists with different backgrounds.

Your absolute and principled honesty is also illustrated by the way in which you give credit to other researchers. Your articles’ and books’ bibliographies are practically exhaustive, you consistently refer to a large number of previous works on the topic, whether they be recent or older publications, in bigger or smaller journals, written by people from your inner circle or by researchers with a completely different background, whether you agree with the analysis or not. You take everyone seriously and you treat people equally, regardless of whether they are absolute beginners or well-established names in the field.

From a more theoretical perspective, I also really like the way in which you manage to undermine long-standing, rusty hypotheses in the linguistic literature by means of a whole number of crystal-clear descriptive and theoretical arguments. These hypotheses cover a range of subjects, such as the position of
left-peripheral topics in English and Romance languages and the syntactic position of the focal element of cleft sentences, to mention but a few. Once you have shed your light on some phenomenon, people no longer write “There is no consensus on...”: the arguments you present are usually so convincing and thought-through that a discussion about that phenomenon is simply no longer possible.

Your research has inspired many other linguists: your work always provides food for thought, it makes the reader want to delve into the matter, to check it in a different language or dialect. One of the great merits of your work is that it rounds off things, but at the same time it creates new starting points for future research. For instance, I am a huge fan of your work on clefts, in which you indisputably show that the clefted element is not located in the left periphery. One issue linguists don’t have to worry about anymore. Other analyses you have conducted and which have largely influenced my own work on French are your many papers on adverbial clauses and main clause phenomena, in which you give the functional linguistic literature the credits it deserves in this area, and in which you demonstrate connections with locality and relativized minimality. I also very much appreciate the way you use notions related to information structure in your syntactic work: with great caution and without reducing them entirely to syntactic features. I haven’t put your ideas about conditional clauses into practice yet, but that’s bound to happen soon. You keep inspiring me.

You also keep finding your own inspiration: you are immensely driven and your output is almost extraterrestrial or at least superhuman (forgive me for these adjectives that may make you frown, given your sober nature). In Leuven there’s this rumor (I’ve told you about it before) that they once generated a scatterplot of the scientific output of Flemish linguists and that you were an absolute outlier. In this sense, the adjectives extraterrestrial and superhuman, which I used above, do have some iconic value . . .

I was first introduced to your work when I read Introduction to Government and Binding Theory, the book with the grey cover that I almost knew by heart, a manual of unprecedented clarity, extremely attractive to a beginning and hungry linguist such as myself at the time. With this book, you have convinced a whole generation of young linguists of the importance of generative linguistics, you’ve explained in a crystal-clear manner what the main empirical questions of the model are, what its explicative power is, what its basic methodology and premises are. On behalf of all these people: thank you!

It’s also fantastic how you were able to bring generative linguistics back to Belgium. I remember very clearly how we heard about your nomination as Odys-
seus professor in Gent about 10 years ago. It was as though generative linguistics in Flanders suddenly had a future again: we, researchers who had thus far been more or less “in hiding” or had sought refuge abroad, could all of a sudden be generative linguists in Flanders out in the open... or at least that’s how it felt. With your appointment, Flanders suddenly acknowledged generative linguistics again; there was activity again, there were conferences, contacts and possibilities, and quite a number of young people were compelled to do research about fascinating syntactic subjects again. Thanks for that as well!

You’ve put West-Flemish on the linguistic map: your papers on complementizer agreement, negative doubling, possessor relations and pronouns in West-Flemish are pearls that – should anyone still have doubts about this – show how apparently small observations on a micro-level can have implications for linguistic comparison on a macro-level and for linguistic theory in general. Thanks to you, “West-Flemish” has made a name for itself in generative linguistics: it has happened so often at conference dinners that I was asked about my mother tongue, and without further explanation, I could answer: “Liliane’s dialect”.

Thank you, Liliane, for participating in scientific committees with your care. It’s a great comfort for many people, generativists and non-generativists alike, that such a principled person, someone who is that objective and honest, watches over the submitted applications.

Thank you for the chances you have given me, for instance the opportunity to speak at the *Subjects Workshop* in 2012. Thanks also for the short chats we’ve had every now and then, usually at some lecture, doctoral defense or conference, sometimes on the phone. It struck me more than once how genuinely offended you still are about certain situations, without shrugging your shoulders about it after some time.

It’s very hard to image you retiring, and I still secretly think (and hope) that you will simply remain as productive as before, because you will be able to dedicate all our time to the core, scientific linguistic research, without the administrative burden that also comes with an academic career. If that’s what you want, I hope with all my heart you will achieve it!

But I also hope that you will have (and take!) a lot more time for all the non-linguistic things you like to do, everything that brings you peace and quiet and warmth, all the small things that make life so beautiful, from frozen yoghurt to cats snoring to intense moments with your loved ones.

Thank you for everything! *Toet in’ droai!* ‘See you later’!
Morphology Driven by Syntax: XP Movement and Allomorphy

Tom Leu*

1 Introduction

In the introduction to syntax class I took with Liliane Haegeman in 1994 in Geneva, I sure learned that syntax was driven by morphology. This was primarily referring to head movement (Pollock 1989), as I recall. Later that decade coming back from the GLOW meeting in Berlin, I had the pleasure of bringing Liliane a handout by Matt Pearson on X(P)-movement which she seemed interested in. Now, two decades later, the relation between syntax and morphology is still on many a syntactician’s daily mind, and (even) the head vs XP-movement issue doesn’t seem to have been decisively settled on all battle grounds. Therefore I will address, in this short paper, how some aspect of morphology is driven by syntax, and more concretely, how an XP movement analysis of verb movement makes more interesting predictions for allomorphy than a head movement analysis, exemplified for one particular kind of case. I do so by adding to Merchant’s (2015) discussion of Greek verb stem allomorphy an analytical option which was not considered by Merchant. I will show how my proposal is able to capture the fact that although four feature (value)s are among the conditioning factors, the resultant allomorphy patterns are systematically limited to a maximally three-way variation.

*I’m very grateful to Arhonto Terzi, Jason Merchant, Heather Newell, and Lena Baunaz for helpful feedback on an earlier version of this paper.
2 Merchant (2015): Spanning allomorphy

Merchant discusses locality constraints on the conditioning context for suppletive allomorphy as exemplified by the voice-aspect system of Greek verbal morphology. The backdrop to his discussion is work by Bobaljik (2000, 2012) and Embick (2010) from which Merchant distills the following proposals:

(1)  
   a. Lexical insertion proceeds bottom-up / root-outward.
   b. Contextual allomorphy requires linear adjacency.

Merchant calls (1b) the Node Adjacency Hypothesis. The assumed syntactic (2a) and corresponding synthetic (2b) structures for the clause and verb, respectively, express the semantic composition of Aspect with a previously formed VoiceP.

\[
\begin{array}{ll}
(2) & \\
   a. & \text{Tense} \\
   & \text{Aspect} \\
   & \text{Voice} \\
   & \text{v} \\
   & \text{VP} \\
   & \text{b.} \\
   & \text{Tense} \\
   & \text{Aspect} \\
   & \text{Voice} \\
   & \text{v} \\
   & \text{Verb} \\
   & \text{v} \\
\end{array}
\]

2.1 Merchant’s explicit problem

Modern Greek verb morphology, as discussed by Merchant, exhibits stem allomorphy sensitive to the combination of Voice and Aspect specifications, which, in some relevant cases, are individually realized/lexicalized, as illustrated in (4d) below. Of particular (though not exclusive) gravity are three verbs which exhibit a three-way stem suppletion pattern that is sensitive to a combination of (the features of) Voice and Aspect.\(^1\) (3) gives the three stems (with the non.active suffix -θ-) for the three suppletive verbs with a description of the relevant conditioning environments (Merchant 2015: p.277, 11).

\[\text{\textsuperscript{1}For fuller discussion, obviously, see the sources cited, among others.}\]

\[\text{\textsuperscript{2}A larger set of verbs raises the same locality issue for non-suppletive allomorphy, all of which also show at most a three-way allomorphy pattern, cf. (Merchant 2015: 281).}\]
The categories that determine the context for allomorph selection are sometimes not overtly realized. This fact has some systematicity to it, to which I will return.

I will, for the sake of argument, assume Merchant’s morphological segmentation to be correct. Since Greek (4d) has overt exponents in both (non-active) Voice and (perfective) Aspect, pruning (i.e. structural annihilation of heads) is not an option. And since Voice and Aspect are individually targeted by VI (Vocabulary Insertion) in (4d), they clearly are not fused.

In [+active] Voice, vocabulary insertion needs to know perfective from imperfective contexts, (4a) vs (4c). And in [-active] Voice, vocabulary insertion, too, needs to know perfective from imperfective contexts, (4b) vs (4d). Hence, assuming (2b), there is a locality problem: Allomorph selection for the stem is sensitive to features of a non-adjacent head. In Merchant’s words: “The form of the stem is determined by the aspect of the verb, but by hypothesis, this aspectual node is not adjacent to the stem in the non-active” (p.281).

The vocabulary items (lexical insertion rules) in (5) capture the right distribution. As Merchant points out (p.280), making stem-allomorph selection sensitive to the presence of /θ/ (which would correctly register non-active perfective) would violate inside-out lexicalization, and is, therefore, on standard assump-

(3) ‘eat’ ‘see’ ‘say’
   a. tro(ɣ)- vlep- le(ɣ)- IMPERFECTIVE
   b. fa(ɣ)- ø- p- ACTIVE.PERFECTIVE
   c. faɣó-θ- iðo-θ- lex-θ-/ipo-θ- NON.ACTIVE.PERFECTIVE

The categories that determine the context for allomorph selection are sometimes not overtly realized. This fact has some systematicity to it, to which I will return.

I will, for the sake of argument, assume Merchant’s morphological segmentation to be correct. Since Greek (4d) has overt exponents in both (non-active) Voice and (perfective) Aspect, pruning (i.e. structural annihilation of heads) is not an option. And since Voice and Aspect are individually targeted by VI (Vocabulary Insertion) in (4d), they clearly are not fused.

In [+active] Voice, vocabulary insertion needs to know perfective from imperfective contexts, (4a) vs (4c). And in [-active] Voice, vocabulary insertion, too, needs to know perfective from imperfective contexts, (4b) vs (4d). Hence, assuming (2b), there is a locality problem: Allomorph selection for the stem is sensitive to features of a non-adjacent head. In Merchant’s words: “The form of the stem is determined by the aspect of the verb, but by hypothesis, this aspectual node is not adjacent to the stem in the non-active” (p.281).

The vocabulary items (lexical insertion rules) in (5) capture the right distribution. As Merchant points out (p.280), making stem-allomorph selection sensitive to the presence of /θ/ (which would correctly register non-active perfective) would violate inside-out lexicalization, and is, therefore, on standard assump-
tions not available.

(5)   a. \( \sqrt{EAT} \rightarrow fa(\gamma) / \quad \text{Voice}[+act] \) Aspect[+perf]

   b. \( \sqrt{EAT} \rightarrow fa\gamma o \quad \text{Voice}[-act] \) Aspect[+perf]

   c. \( \sqrt{EAT} \rightarrow tro(\gamma) \quad (\text{elsewhere form, found in} \) [\pm \text{active}].[\text{imperfective}])

Of particular relevance to the proposal I will develop in section 3 is that in the imperfective, (5c) is used irrespective of the Voice specification. Hence it is the perfective Aspect that correlates with the more specific allomorphy, where allomorph allomorph selection differentiates Voice specification.

2.2 Merchant's proposal

Keeping to the structure in (2b), Merchant loosens the locality requirements on allomorphy to a sufficient degree by calling upon the notion of span (Abels & Muriungi 2008, Taraldsen 2010, Svenonius 2012) which has been formulated in some of the nanosyntactic literature. Merchant (p.288) mentions Svenonius' (2012) formulation of a span as "a complement sequence of heads ... in a single extended projection." This shall be sufficiently precise for our purposes. His proposal now consists of two claims, which can be informally rendered as follows.

(6)   a. Only a span can be targeted by vocabulary insertion.

   b. Allomorphy can be conditioned only by an adjacent span.

(6a) is standard in the spanning literature. (6b) is Merchant's replacement of (1b), which he calls the Span Adjacency Hypothesis. Note that whether the heads of a span are lexicalized individually or in a portmanteau fashion is immaterial to their involvement in allomorph-selection in a subjacent node. This is, of course, expected (for inside-out sensitivity) given inside-out lexicalization.

2.3 Merchant's (implicit) prediction

Of interest to my contribution here is the prediction, implied in Merchant's proposal, regarding the variety of allomorphy patterns allowed and hence expected.

---

3 This allows an X-bar friendly variant of aspects of Brody's (2000) mirror theory proposal.

4 "Allomorphy is [...] conditioned locally [...] by features in adjacent spans, whether or not those spans are themselves lexicalized by Vocabulary items." (Merchant 2015: 294)
In the empirical domain of Modern Greek verbs considered by Merchant, we observe three allomorphy patterns: (a) lack of allomorphic variation (e.g. enθarin- ‘encourage’, p.283, 20, 15); (b) two-way allomorphy (e.g. din- / di- ‘dress’, p.283, 20, 12); and (c) three-way allomorphy (e.g. empne- / empnef- / empnefs- ‘inspire’, p.283, 20, 3). The latter pattern constituted the crucial cases for his proposal.

Interestingly, though, Merchant’s proposal would also allow a four-way allomorphy pattern in which different stem variants would be used in active.imperfective vs. nonactive.imperfective. This, however, does not seem to be attested in Modern Greek. In other words, his proposal may be too permissive. More specifically, not only is there maximally three-way stem allomorphy, but all such cases exhibit the same conditionning contexts. Furthermore the conditioning contexts in the two-way allomorphy cases forms a subset of the three-way pattern. In particular, there is no case of sensitivity to [-active] in the imperfective. On Merchant’s account, there is no reason to expect this asymmetry.

In the next section, I will briefly outline an alternative which allows only up to a three-way allomorphy pattern, excluding the unattested fourth case.

3 An XP-alternative

Key to the limitation to a three-way allomorphy pattern is, I believe, the observation that, while indeed stem selection shows sensitivity to both Aspect (perfective vs imperfective) and Voice (active vs nonactive), the Voice-sensitivity is limited to one of the values of the aspectual contrast. Concretely, stem selection is sensitive to the [-±active] Voice distinction only in the perfective. On Merchant’s analysis this is accidental. I will try to provide a rationale for it.

I will, without discussion, adopt Merchant’s (and hence indirectly Rivero’s 1990) data structural lay of the land, in particular, the syntactic hierarchy in (2), the treatment of perfective and imperfective as the two values of the same syntactic head Aspect6, and the overt morphological segmentation (i.e. the proposal that -ik- spells out [perfective] in [-active] [+past] environments, and that -θ- spells out [-active] in [perfective] environments7).

5Christopoulos & Petrosino (2017) propose another account of the Merchant facts, respecting strict adjacency by making use of post-syntactic re-bracketing. Their proposal is, however, inert relative to the limitation to the three-way allomorphy pattern I observe.

6Abandoning this assumption might allow for a more elegant variant of my alternative, but it would also require a more spacious elaboration.

7These relevant contexts are all linearly adjacent to the lexicalisation target on my proposal,
What I crucially reject from his proposal is the idea that the inflected verb corresponds to a complex syntactic head derived by head-movement. Instead I assume verb movement to be XP-movement (Koopman & Szabolcsi 2000, Mahajan 2003). More concretely, the verb moves qua vP or VoiceP (or...), i.e. in a successive-cyclic or a roll-up fashion (Cinque 2005).

Let me propose that Modern Greek distinguishes perfective and imperfective in its syntactic derivation: Asp_{perfective} attracts Voice(P), but Asp_{imperfective} attracts vP, rather than Voice.

(7)

\[
\begin{align*}
\text{(a)} & \quad \text{AspP} \\
& \quad \text{VoiceP} \\
& \quad [V-v] \pm \text{active} \quad \text{Asp} \\
& \quad \text{[perf]} \\
& \quad t_{\text{VoiceP}} \\

\text{(b)} & \quad \text{AspP} \\
& \quad vP \\
& \quad V-v \\
& \quad [\text{imperf}] \quad \text{VoiceP} \\
& \quad t_{vP} \quad \pm \text{active} \\
& \quad t_{vP}
\end{align*}
\]

Furthermore, Asp_{imperfective} seems to have the property of hiding from PF everything its mother dominates (i.e. itself and its complement), cf. Merchant (2001) on sluicing.

On this proposal, out of the four logically possible feature combinations (as–ignoring traces.

8 For the present discussion, the crucial difference between XP-movement and head movement emerges on the standard assumption that head movement disallows excorporation. If that assumption is abandoned, my proposal can be mimicked in head movement terms.

9 It may not be accidental that it is the imperfective - aspect which references “the internal temporal structure of the [event]” (Comrie 1976: 24) - that requires a closer syntactic relation with vP, i.e. with the event.

10 An effect being that there is no overt mirror principle violation. The fact that the relevant heads are non-overt in the configuration in which their overtness would violate the mirror (i.e. imperfective everywhere and voice in the imperfective) may well suggest a more principled analysis of this correlation, perhaps in terms of phrasal spellout (cf. Starke 2009, Caha 2009). I will leave this for other (occasion)s.
suming, as Merchant does, a binary Voice and a binary Aspect opposition) exactly three strictly linear stem adjacency possibilities are derivable (8). The linearly adjacent environment of the verb stem is identical in (8c) and (8d):

(8) a. \([V+v] - [\text{Voice}_{\text{active}}] - [\text{Aspect}_{\text{perfective}}]\)  
b. \([V+v] - [\text{Voice}_{\text{non-active}}] - [\text{Aspect}_{\text{perfective}}]\)  
c. \([V+v] - [\text{Aspect}_{\text{imperfective}}] - [\text{Voice}_{\text{active}}]\)  
d. \([V+v] - [\text{Aspect}_{\text{imperfective}}] - [\text{Voice}_{\text{non-active}}]\)

Therefore at most a three-way allomorphy pattern is admissible under strict linear adjacency.

(9) a. \(\sqrt{EAT} \rightarrow fa(\gamma) / \text{Voice}[+\text{act}]\)  
b. \(\sqrt{EAT} \rightarrow fa\gamma o / \text{Voice}[-\text{act}]\)  
c. \(\sqrt{EAT} \rightarrow tro(\gamma) \quad (\text{elsewhere})\)

In the imperfective, the stem is not adjacent to Voice and hence the elsewhere form is used. To enhance clarity, please consult the tree diagrams for the 1.SG.NONACTIVE.PERFECTIVE.PAST form \(fa\chi\delta-\theta-\text{ik}-a\) (8b) = (9b) in (10).

(10) \(fa\chi\delta-\theta-\text{ik}-a \quad (1.SG.NONACTIVE.PERFECTIVE.PAST)\)

and for the 1.SG.NONACTIVE.IMPERFECTIVE.PAST form \(troy-\text{umun}\) (8d) = (9c) in (11).
This analysis makes a second prediction, essentially a flip-side prediction of the first: Given that T is linearly adjacent to Aspect in the perfective, and linearly adjacent to Voice in the imperfective, it seems natural to expect the possibility of allomorphy in T (tense-agreement) to be sensitive to a \([-\text{active}]\) Voice contrast in the imperfective but not in the perfective. This is illustrated here by the \(1\text{SG.PAST}\) forms of the verb \(t\rho\gamma\delta\text{-}o\mu\nu\delta\) ‘I eat’ in (12)–(13), comparing the tense-agreement forms (following Merchant taken as fusional, which may be too simplistic) across the four cells of the \(\pm\text{active}\) and \(\pm\text{perfective}\) dimension.

(12) Imperfective (1SG.PAST, ‘eat’)

a. ACTIVE: \(\acute{\text{e}}\text{-t}\rho\gamma\delta\text{-}\alpha\)

b. NON-ACTIVE: \(t\rho\gamma\delta\text{-}o\mu\nu\delta\)

(13) Perfective (1SG.PAST, ‘eat’)

a. ACTIVE: \(\acute{\text{e}}\text{-f}\acute{\text{a}}\rho\gamma\delta\text{-}\alpha\)

b. NON-ACTIVE: \(f\acute{\text{a}}\rho\gamma\delta\text{-}\theta\text{-ik}\text{-}\alpha\)

The non-active imperfective \((12b)\) is the odd one out. Since only in the imperfective is T linearly adjacent to Voice on my proposal (cf. (11)), this can be suppletion conditioned by \([-\text{active}]\) Voice, assuming that the morphosyntactic features stick around for another while after VI. The pattern seems to be rather pervasive in Greek verbal morphology across different conjugation classes and different
Whether this prediction withstands more detailed scrutiny, only such scrutiny will be able to tell.

4 Conclusion

Modern Greek verb stem allomorphy shows sensitivity to perfective vs imperfective Aspect, and within perfective Aspect to the [±active] Voice distinction. On the assumption that Verb, Voice, and Aspect are individual syntactic heads in that (hierarchical) order, the sensitivity to both Voice and Aspect has been taken argued, by Merchant, to show the need to loosen the locality restrictions on the conditioning of vocabulary insertion from strict linear node adjacency to span adjacency.

Observing that the observable allomorphy patterns do not exploit the full potential of diversity allowed under Merchant’s approach, but instead systematically fail to distinguish [±active] Voice in the imperfective, I propose an analysis in which the verb stem is linearly adjacent to Voice only in the perfective, and is adjacent to Aspect in the imperfective. On this analysis, strict linear adjacency in allomorph selection without reference to spans can be maintained.

The contrast in linear adjacency is derived on the assumption that word formation is accomplished syntactically, by XP-movement, in conjunction with the proposal that Voice attracts vP, perfective Aspect attracts VoiceP, and, crucially, imperfective Aspect attracts vP.

The proposal may catch a second fly on the same boat, providing the beginning of a principled approach to the generalization that the tense-agreement exponent shows sensitivity to nonactive in the imperfective but not in the perfective.

References


---

I’m grateful to Arhonto Terzi for very helpful confirmation of this prediction.


Pearson, M. 1999. X(P)-movement and word order typology. ‘direct’ vs. ‘inverse’ languages. GLOW handout.


Multilingualism as the new comparative syntax

Terje Lohndal

1 A model of a linguist

Liliane Haegeman has made groundbreaking and lasting contributions to the study of grammar, and it is a great pleasure to be able to honor such a wonderful linguist and friend by way of a small contribution in the present festschrift. Her work is characterized by its focus on argumentation, rigor, and theoretical perspicuousness. During her career, she has covered a range of areas in addition to ‘core’ syntax, as also emphasized in another festschrift (Aboh et al. 2017: 3): Dialect variation, register variation, first and second language acquisition. Her work on comparative syntax can be characterized as she herself characterizes the field, where emphasis should be put on the very last sentence in the quote.

An important development in generative syntax over the last fifteen years is the revival of the interest in the comparative study of language. In fact, a major criticism leveled against early generative grammar concerned its central interest for conceptual problems often to the detriment of the empirical study of language. When empirical data were considered, this was often only to serve a theoretical point, and in the analyses a small range of data of only a handful of languages was taken into account, standard English occupying a central position among the languages under examination. Over the past twenty years, we have witnessed a surge of comparative work along various dimensions. In this new comparative syntax, careful study of empirical data takes a central position with a
stimulating two-way interaction between theoretical developments and empirical study. (Haegeman 1997: 1)

This quote highlights the importance of a close interplay between description and theory as they both depend on each other. This interdependence has always been crystal clear in Liliane’s work as she has worked to identify the linguistic properties that can vary across languages and those that are constant. This puts her work firmly into the generative tradition trying to unearth what knowledge of language is and how this knowledge can be acquired. As she continues in her already quoted chapter, “The question of acquisition focuses on the issue of how much of our linguistic knowledge is due to experience, the linguistic input provided by the environment we are exposed to, and how much is due to a predetermined mental faculty” (Haegeman 1997: 1).

Methodologically, her work has made use of several different methods. Since her earliest work (Haegeman 1983), she has been a keen consumer of naturally occurring data, which she has either accessed through corpora or gathered herself. She has of course also relied heavily on acceptability judgments both in investigating her native West Flemish but also in her investigations of English.

Liliane’s work has always highlighted how linguistic competence needs to be investigated broadly, and how different sources of data speak to different aspects of this competence. And she has set a high standard for the tight relationship between empirical generalizations and theoretical analysis. With that in mind, the next section will try to say something about one aspect of current developments, which in many ways can be viewed as an integral part of the work on comparative syntax that the quote from Haegeman (1997) highlights.

2 The formal grammar of multilingualism

Comparative syntax has been the zeitgeist of formal grammar since the early 1980’s. In recent years, we may say that a new zeitgeist has emerged. It is not the only zeitgeist, but it is one of several, and it is the formal study of multilingual data. Traditionally, multilingualism has not been at the core of formal generative studies. This is among others made clear in the following famous quote:

Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such
grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance (Chomsky 1965: 3).

It was always clear that this was an idealization - no one thought that such an ideal speaker-listener or such a speech-community actually existed. However, it was a very useful idealization as it made it easier to develop theories of complex empirical phenomena. It enabled the grammarian to abstract away from issues that would make it hard to extract generalizations and then start to develop descriptively adequate grammar fragments.

It took some time, but eventually work on second language acquisition was established within a generative frame. Whether the logical problem of language acquisition applied to the acquisition of an additional language in adulthood, was one of the main goals of the early research in this area, see e.g., Clahsen & Muysken (1986), Flynn (1987), Schwartz (1987), Bley-Vroman (1989), White (1989), Schachter (1990). Put differently, do language-specific constraints reduce the hypothesis space in learning a second language, similarly to what is argued to be the case for first language acquisition? Textbooks such as Hawkins (2001), White (2003) and Slabakova (2016) document the success of this approach, and (Rothman & Slabakova 2017) is a state-of-the-art paper outlining the changes in generative approaches to second language acquisition over the years and connecting them to the current prominent approaches and trends. Recently, third language acquisition has also become a productive area of investigation within formal models, cf. e.g., Rothman (2011, 2015) and Westergaard et al. (2016) and references therein for discussions and comparisons of different models.

Another area which departed from the idealization, focused on whether or not there are formal constraints on code-switching or language mixing (e.g., Pfaff (1979), Poplack (1980), Sankoff & Poplack (1981), Woolford (1983), Di Sciullo et al. (1986), Belazi et al. (1994), MacSwan (1999, 2000, 2005), Muysken (2000, 2015), Myers-Scotton (2002), van Gelderen & MacSwan (2008), González-Vilbazo & López (2011, 2012), see MacSwan (2014) for a review and see Riksem (2018) for additional discussion). Studying language mixing may provide us with a typology of what elements are possible to mix and which are resistant (to various degrees) to mixing across languages. The following quote from Riksem (2018: 43) emphasizes this point:
In general, language mixing may constitute a potential window into our language capacity; the conditions and restrictions on language mixing can tell us which linguistic elements are possible to mix, and whether some are more available or resistant to mixing than others. Thus, studies of language mixing may refine and deepen our understanding of grammatical theory (Muysken 2000, Gardner-Chloros 2009, González-Vilbazo et al. 2013).

By way of illustration, let us consider mixing between English and Norwegian. I will use examples from the heritage language American Norwegian spoken in the USA (see Haugen (1953) for an extensive presentation of the language and its background). The specific examples here are drawn from the Corpus of American Nordic Speech (CANS; Johannessen (2015)). Consider (1) first.

(1) a. Jeg teach-a #ález grade-en I teach-PAST # first # grade-DEF.M.SG ‘I taught the first grade.’

b. Så kan du mow-e litt lawn then can you mow-INF some lawn-UNDEF.SG ‘Then you can mow some lawn.’ (coon_valley_WI_07gk)

These examples illustrate a typological generalization (see also Afarli (2015) and Riksem et al. (In press) on American Norwegian): One of the language is the language providing the grammatical structure, whereas lexical items can come from either language. This can be seen clearly in (1b), where the structure of the sentence exhibits Verb Second (V2; see Eide & Hjelde (2015) and Westergaard & Lohndal (2018) for more on V2 in American Norwegian). There are also additional examples showing that English lexical items can appear in structures that are clearly Norwegian.

(2) a. Å celebrat[e]-e birthday-en hennes to celebrat-INF birthday-DEF.M.SG her ‘To celebrate her birthday.’ (coon_valley_WI_06gm)

b. etter middag-en # vi # satt på deck-en hans after dinner-DEF.M.SG # we # sat on deck-DEF.M.SG his ‘After dinner, we sat on his deck.’ (westby_WI_01gm)

In both of the examples in (2), the speakers are clearly using a Norwegian struc-
ture for possessives: The possessive is post-nominal, and the nominal, even if
the lexical item is English, has the Norwegian definiteness marker. A syntactic
theory needs to be able to account for this, and a theory assuming that the syn-
tactic structure is based on the lexical items itself would not be a descriptively
adequate theory as it would predict that the structure should be English, not
Norwegian. Space does not allow me to elaborate on what such a theory could
look like, the reader can consult Grimstad et al. (2014), Riksem (2018), Grimstad
et al. (in press) and Riksem et al. (in press) for one alternative.

3 Conclusion

Work on second language acquisition and aspects of language mixing shows that
formal approaches also have a lot to contribute when it comes to multilingual
data. In many ways, they complement more traditional work focusing on vari-
atation between varieties, but instead of looking at variation between individuals,
they look at variation within individuals. They help us to create better models of
our language competence and its scope. As such, they may constitute the new
comparative syntax in the years to come, trying to follow the very high standards
set by Liliane Haegeman and others.

References

Aboh, E., Haeberli, E., Puskás, G. & Schönengerber, M. (eds.). 2017. Elements of
Comparative Syntax: Theory and Description. Berlin: Mouton de Gruyter.
Åfarli, T.A. 2015. A Syntactic Model for the Analysis of Language Mixing Phe-
nomena: American Norwegian and Beyond. In B.R. Page & M.T. Putnam
(eds.), Moribound Germanic Heritage Languages in North America: Theoretical
Bley-Vroman, R. 1989. What is the logical problem for foreign language learn-
ing? Linguistic Perspectives on Second Language Acquisition 4. 1–68.
Clahsen, H. & Muysken, P. 1986. The availability of universal grammar to adult
and child learners: a study of the acquisition of German word order. Second
Language Research 2. 93–119.


176


To Liliane, with appreciation

Shigeru Miyagawa

I first met Liliane in Japan when she came to give a talk at Kanda University of International Studies. With Memo Cinque and others, we toured Tokyo together and quickly became friends. And the friendship blossomed into professional collaboration, including co-organizing a LSA panel on MCP and jointly giving a plenary talk at GLOW.

Around the time that we met, I had completed the manuscript for Why Agree? Why Move?, which was published as a Linguistic Inquiry monograph in 2010. I proposed Strong Uniformity, an instantiation of Chomsky’s (2001) Uniformity Principle; SU states that every language shares the same set of grammatical features, and each language manifests these features in some fashion. When one makes a sweeping proposal of this sort, one quickly digs holes to fall into; a sign of a promising theory is that you can dig out of at least some of them.

A particularly embarrassing hole that I dug for myself was the prediction that Strong Uniformity made that a language such as Japanese has ϕ-feature agreement at C. This is embarrassing because Japanese is known as a prototypical agreementless language. Everyone knows this! But predictions are predictions—I either show that it is borne out despite all appearance to the contrary, or abandon the project. In attempting to address this problem, Liliane’s work on the MCP came into sharp focus. I was trying to argue that the politeness marking in Japanese (−mas-) is second-person agreement at C, because it “agrees” with the hearer, and it occurs at C [Miyagawa (1987)]. The politeness marking is an MCP. But it is one thing to stipulate that the politeness marking is ϕ-feature agreement, something else altogether to show that it is a credible way to view it as such. The big breakthrough came when Liliane invited me to give a presentation at the 2010 Ghent workshop on clause-typing and main clause phenomena. In searching for a topic to discuss, I discovered data from Basque (courtesy of Karlos Arregi) that gave credence to the idea that the politeness marking in Japanese is 2nd person agreement, and it is at C.

The so-called allocutive agreement in certain Basque dialects agrees with the
hearer, thus it is always 2nd person, despite the fact that there is no 2nd person entity in any of the argument positions. Oyharçabal (1993). The allocutive is a regular form of agreement, as shown by the fact that it competes for position with the “regular” 2nd person agreement. Its function is politeness (formal/colloquial), and it occurs at C. In fact, Oyharçabal, who wrote the article on allocutive agreement, refers to an earlier article of mine (Miyagawa 1987) and observes that the allocutive agreement in Basque has essentially the same distribution as the politeness marking in Japanese. I remember that Liliane was quite excited about the data, which gave me confidence that I’m onto something. I got to publish this work in the collection of papers from the workshop (Main Clause Phenomena: New Horizons, 2012) that was co-edited by Liliane. The phenomenon of allocutive agreement is novel in generative grammar, and a number of linguists picked it up after the publication of the book and produced interesting studies using a variety of languages. One of them was Vera Zu, who completed a Ph.D. dissertation on the topic in 2017 that substantially extended the idea of allocutive agreement beyond Basque and Japanese. Encouraged by the discovery of allocutive agreement, I myself went on to write a second LI monograph, Agreement beyond phi, which was published in 2017.

Liliane and I collaborated on two big projects, both stemming from our mutual interest in main clause phenomena. We co-organized a Linguistic Society of America panel in 2013 which was held in Boston, and invited Jim Huang, Luigi Rizzi, and Raffaella Zanuttini to join us. In 2016, we gave a joint plenary talk at GLOW 39 held at the University of Göttingen. In that talk, Liliane went over some of the major achievements of her enormous project on main clause phenomena dating back to the early 2000’s, while I introduced the notion of allocutive agreement and its implications for linguistic theory. It was one of the most enjoyable talks I’ve ever given, and certainly one that I learned a great deal from thanks to Liliane.

I cherish my friendship with Liliane, and I am deeply grateful to her for bringing me into her world of linguistic study that she helped to pioneer. It has certainly been rewarding for me, not the least of which is because I got to work with her.

References

Oyharçabal, B.n. 1993. Verb agreement with non arguments: On allocutive

Hoe cool is dat wel niet!
How pseudo-questions and expletive negation in Dutch

Rachel Nye and Lieven Danckaert*

1 Introduction

This paper focuses on the Dutch equivalents to structures such as (1), referred to by Nye (2009, 2011) as how-pseudo questions (HPQs). On the surface, HPQs resemble how degree questions (HDQs) (2), but combine the surface word order of a matrix question with an interpretation more usually associated with exclamative structures.

(1) How cool is that! HPQ
(2) How old is he? HDQ

As shown in (3) and (4), very similar structures are also available in Dutch:

(3) Hoe cool is dat! HPQ
    how cool is that

*There is much that we could – and would like to – say in tribute to Liliane, both in terms of her contribution to the field of linguistics, and regarding the support that she has provided us both with. Taking the saying “actions speak louder than words” to heart, a revival for this particular occasion of the former 'PhD student + Postdoc buddy' team established by Liliane during the early GIST days seemed the most suitable way to convey these sentiments. The connections this paper shows to Liliane’s work, in terms of the languages studied, the linguistic topics touched upon and the approach taken to the data should be apparent throughout.

Throughout this paper, in constructed examples or examples from a spoken language source, we punctuate HDQs with a question mark and HPQs with an exclamation mark. In attested written examples, there is considerable variability in how HPQs are punctuated (as discussed by Nye (2009) for English), and the original punctuation is maintained when such examples are reproduced. Punctuation proves an equally poor indication for determining the HPQ vs. HDQ status of a given string in Dutch as it does in English.
How cool is that!

(4) Hoe oud is hij? HDQ
    how old  is he
    'How old is he?'

HPQs are common in contemporary English and Dutch, and they differ from what one could call 'standard exclamatives', such as English (5) and Dutch (6), in two important respects. First, HPQs are more colloquial, whereas the exclamatives in (5) and (6) are primarily associated with a more formal register. Second, only HPQs have a string-identical interrogative counterpart.

(5) a. How tall that building is!
    b. What a tall building that is!

(6) a. Wat is dat gebouw hoog!
    what is that building high
    b. Wat is dat een hoog gebouw!
    what is that a  high  building
    'How tall that building is!'

Our first aim is to offer a description of the Dutch HPQ pattern exemplified in (3). Concretely, we show that Dutch and English HPQs have very similar properties: in line with Nye (2009, 2011) we argue that despite the apparent surface resemblances, HPQs do in fact differ structurally from HDQs, in Dutch as well as in English. On the basis of this we conclude that the interpretive properties of HPQs cannot be explained as a pragmatic effect arising from the use of a HDQ in a particular context. Rather, the difference in interpretation of HPQs and HDQs results from a difference in the underlying syntax of these two structures. As a result, despite the presence of subject-auxiliary inversion in both (1) and (2), HPQs are best categorised as a type of exclamative, rather than as interrogatives.

Secondly, on the basis of the Dutch data we develop a new argument for the claim that HPQs should be distinguished structurally from HDQs, which goes beyond those already put forward by Nye (2009). In particular, we demonstrate that in Dutch HPQs it is possible for the sentential negator niet 'not' to occur, without any of the negative force with which it is typically associated. The fact that this type of ‘expletive negation’ is not available in Dutch HDQs again suggests that HPQs and HDQs do not have the same syntactic structure, despite typically having the same word order.

Section 2 of this paper briefly recapitulates the main arguments put forward
by Nye (2009, 2011) on the basis of English HPQs for the claim that HPQs are structurally distinct from HDQs, and applies these to Dutch HPQs. In section 3, the phenomenon of expletive negation in Dutch HPQs is discussed and in section 4, HPQs and other inverted exclamatives are compared to rhetorical questions. Section 5 suggests directions for future research and section 6 concludes.

2 Distinguishing HPQs from HDQs

Nye (2009) provides evidence to show that although HPQs can be string-identical to HDQs, HPQs differ from HDQs both in interpretation and, in some cases, in form. Three key properties identified by Nye (2009) as distinguishing HPQs from HDQs are applied to the Dutch data, with the same results shown to hold for Dutch as for English.

2.1 Questions

First, HDQs introduce a question into the discourse, whereas HPQs do not. For example, a HDQ can felicitously be used in the exchange in (7), where A’s utterance is a genuine information-seeking question, to which the B’s reply constitutes a pragmatically appropriate answer.

(7) A: Signalen worden door de satelliet verstuurd met de snelheid van het licht. Hoe snel is dit (uitgedrukt in km/sec)?

B: 300.000 km per seconde. 300,000 kilometres per second ‘300,000 kilometres per second.’

In contrast, as shown in (8), HPQs are incompatible with an answer that provides a piece of new information. This is of course not to say that they cannot

---

Last accessed 30.06.2018.
be replied to at all. A response indicating agreement with the speaker’s assessment, such as the one in (8B’), is typically felicitous.

(8)  

A:  Wow hoe snel is dit?  

wow how fast is this  

‘Wow, how fast is this!’

B:  #Wel, het heeft precies 2 uur en 5 minuten geduurd.  

well it has precisely 2 hours and 5 minutes lasted  

‘Well, it took exactly 2 hours and 5 minutes.’

B’:  Ja, inderdaad, dat was echt snel.  

yes indeed that was really fast  

‘Yes indeed, that was really fast.’

We can conclude that while HDQs have interrogative illocutionary force, HPQs do not. However, on a par with bona fide exclamatives (cf. Portner & Zanuttini 2000), nor do HPQs have the same distribution as declaratives, as attested to by the fact that they cannot felicitously be used to answer a question, a point exemplified by the dialogue in (9):

(9)  

A:  Werd het pakket snel afgeleverd?  

became the parcel quickly delivered  

‘Was the parcel delivered quickly?’

B:  #Hoe snel was dat!  

how fast was that  

‘How fast was that!’

B:  #Wat was dat snel!  

what was that fast  

‘How fast that was!’

C:  (Ja,) de levering was echt snel.  

yes the delivery was really fast  

‘(Yes,) the delivery was really fast.’

This suggests that the HPQ pattern is in fact a type of exclamative.

2.2 Evaluativity

Another property shared by genuine wh-exclamatives and HPQs is the fact that both show polarity-insensitive evaluativity, which HDQs do not. The concept of evaluativity (in the realm of adjectival constructions) is defined in Rett (2015:1) as follows: “[a]n adjectival construction is evaluative iff it makes reference to a degree which exceeds a contextually valued standard”. Adjectival predicates in HPQs do seem to qualify as evaluative expressions: whereas HDQs ask about the extent to which the property expressed by the adjective holds, HPQs exclaim about the fact that this property holds to a great extent. For example, in the example in (10) it is presupposed that the Herengracht is beautiful; what is asserted is that it is indeed very beautiful, prettier than other parts of Amsterdam at the same time of year, for instance.

(10) Maar kijk eens aan: hoe mooi is de Herengracht in de zomermaanden?

‘But just look at that: how beautiful is the Herengracht during the summer months!’

In contrast, no such effects are present in a regular HDQ such as (11) (which is the title of a government brochure about spatial planning). Not only does this question lack the presupposition that The Netherlands will be beautiful in the near future, given the interrogative semantics there is also no assertion that the country will score high on a scale of beauty.

(11) Hoe mooi is Nederland morgen?

‘How beautiful will The Netherlands be tomorrow?’

Following Rett (2011), we can conclude that Dutch HPQs behave like structures whose status as exclamatives is not debated in relation to the property of evaluativity.

2.3 Intensifiers

Third, as observed in Nye (2009), HPQs – like exclamatives and in contrast to HDQs – can contain intensifiers such as totally and incredibly which modify the adjective in the how-phrase. The same holds for Dutch. Intensifiers such as bangelijk ‘scarily’, kei- ‘very’ (lit. ‘boulder’), verdomd ‘damn’, vet ‘fat’, vreselijk ‘terribly’ and fucking (borrowed from English) can modify the adjective within the hoe-phrase in Dutch HPQs. Crucially, this is also the case for exclamatives, in contrast to HDQs, where such intensifiers are typically excluded.

(12) a. WTF hoe vet cool is da!!!
   WTF how fat cool is that
   ‘WTF, how very cool is that’

b. Die blauwe op 1’50; hoe bangelijk goed ziet die der uit zeg!!
   that blue at 1:50 how scarily good sees that PRT PRT say
   ‘The blue one at 1:50; how scarily handsome does he look!’

c. Hoe vreselijk erg is het om het WEL te doen?:
   how terribly bad is it to to PRT to do
   ‘How terribly bad is it to actually do it!’

d. En hoe kei cool is je poncho geworden [...]?
   and how very cool is your poncho become
   ‘And how cool has your poncho turned out to be!’

e. Hoe verdomd geil kan je worden van sushi:
   how damn horny can you become of sushi
   ‘How damn horny can you get from sushi!’

f. HOE FUCKING COOL IS GUARDIANS OF THE GALAXY VOL. 2?!?
   how fucking cool is Guardians of the Galaxy vol. 2
   ‘How fucking cool is Guardians of the Galaxy vol. 2?!?’

---

As shown in (13), none of the strings given in (12) can plausibly be coerced into a question (i.e. HDQ) interpretation:

(13)  
   a. *Hoe vet cool is da?  
   b. *Hoe bangelijk goed ziet die der uit?  
   c. *Hoe vreselijk erg is het om het WEL te doen?  
   d. *En hoe keicool is je poncho geworden?  
   e. *Hoe verdomd geil kan je worden van sushi?  
   f. *HOE FUCKING COOL IS GUARDIANS OF THE GALAXY VOL. 2? 

In other words, there are contexts in which HDQs and HPQs are in fact formally non-identical, which strongly suggests that the difference between the two patterns is not a matter of pragmatics alone. In the following section, we discuss an additional phenomenon which also differentiates Dutch HPQs from HDQs on formal grounds.

3  Expletive negation in exclamatives


For English, the observation that exclamatives can permit expletive negation dates back to Jespersen (1924: 323) who states that “in exclamations [...] very often it does not matter whether not is added or not”, providing the example given here as (14) by way of support:

(14)  
   How often have I (not) watched him!

15.01.2018.
However, it is by no means the case that expletive negative is productively available in all present day English exclamatives (Jespersen’s example is characterised by Horn (2010: 123) as “now somewhat quaint-sounding”). In the non-inverted equivalent (15), for instance, not seems to be excluded (at least when it is interpreted as expletive negation).

(15) How often I have (*not) watched him!

The availability of expletive negation in Dutch exclamatives has been noted by Espinal (2000: 66), among others, who gives the following examples (her (32a,b)):

(16) a. Wat heeft hij niet een vragen gesteld!
   what has he not a questions raised
   ‘He raised so many questions!’

   b. Wat heeft hij niet een ellende veroorzaakt!
   what has he not a mess caused
   ‘He created such a mess!’

The examples in (16) do not involve HPQs but rather ‘standard’ wh-exclamatives (see also (6)). Note that the (invariant, but possibly discontinuous) string wat een (sg. or pl.) NP (lit. ‘what a NP’) cannot be used in interrogative contexts, which disambiguates the structures in (16) from interrogatives.13

Let us now turn to negation in HPQs. For English, Nye (2009: 18) provides the attested example presented here as (17). To the extent that such marginal cases are accepted, not has the expected force of sentential negation, negating the propositional content: (17) is an exclamation about the lack of vigilance. The interpretation and acceptability of negation in canonical non-inverted exclamatives seems to be similar (18). Sentential negation not seems even more marginal in HDQs (19).

(17) ?How vigilant are they not!

12 Inverted exclamatives such as (14) are differentiated from HPQs in terms of both prosody and register: inverted exclamatives such as (14) seem out-dated, formal and literary, much as the noninverted equivalent in (15) does, while HPQs, in contrast, are characteristic of contemporary, informal, colloquial speech. The possibility of expletive negation in inverted exclamatives but not HPQs in English, as discussed in this section, provides another piece of evidence in favour of the view that these structures should be distinguished.

13 As such, although the English translations Espinal (2000) provides for (16a) and (16b) are declaratives, these could perhaps more accurately be rendered as ‘What a lot of questions he raised!’ and ‘What a mess he caused!’ respectively.
Although in many cases Dutch HPQs are string-identical to their English equivalents, one striking difference is the ease with which niet ‘not’ can occur: some speakers in fact have a preference for it to be present in HPQs. What is important to note is that what we are dealing with here is expletive negation, that it is to say ‘the negation is not interpreted according to its canonical logical meaning’ (Delfitto & Fiorin 2014: 284). In fact, only on this interpretation are such cases grammatical. Two attested examples are given in (20):

(20)  
(a) Ik heb geen kind en ik wil ook Disney. Hoe erg is dat niet?\(^4\)  
‘I don’t have a child and I also want Disney. How bad is that!’
(b) Een boek dat zich om laat bouwen tot een flipperkast, hoe leuk is dat niet?\(^5\)  
‘A book which can be transformed into a pinball machine, how cool is that?’

In these examples niet does not negate the truth of the propositional content of the sentence, but rather appears to have an intensifying function (if any). The crucial observation is that niet in Dutch HDQs – to the extent that it is acceptable at all – can only be interpreted as conveying canonical sentential negation, on a par with what can be observed for English.

(21)  
(a) ??Hoe erg is dat niet?  
‘How bad is that not?’
(b) ??Hoe leuk is dat niet?  
‘How nice is that not?’

Observe that Dutch HPQs are fully disambiguated from HDQs on formal grounds

when expletive negation is accompanied by the positive polarity particle wel, with which it can (and readily does) occur in combination. Some attested examples are given in (22):

(22)  

a.  Elektriciteit, hoe belangrijk is dat wel niet\(^{16}\)  
    electricity how important is that PRT not  
    ‘Electricity, how important is that!’

b.  Sinds mijn kindertijd – hoe lang is dat wel niet geleden? –  
    since my childhood how long is that PRT not past  
    hanteer ik even vlot het Nederlands als het West-Vlaams  
    handle I equally fluently the Dutch as the West.Flemish  
    en het castellano [...].\(^3\)  
    and the Castellano  
    ‘Since my childhood – how long ago is that! – I’ve been equally fluent in Dutch, West Flemish and Castellano.’

The elements wel and niet can also occur in combination with a seemingly similar function and interpretation in other exclamative clauses (cf. (23a)), but crucially not in interrogatives with question force (cf. (23b)).

(23)  

a.  Wat er allemaal wel niet gedaan is.\(^{38}\)  
    what EXPL all.ADV PRT not done is  
    ‘How many things have been done!’

b.  *Wat is er allemaal wel niet gedaan?  
    what is EXPL all.ADV PRT not done

To conclude, we now have another argument in favour of the claim that – at least in Dutch – HPQs differ structurally from HDQs. Under purely pragmatic accounts of their differences in interpretation, the differing distribution and interpretation of the negator niet (and the positive polarity particle wel) in the two contexts remains unexplained.

\(^{3}\)http://uilenvlucht.jotemo.be/?m=201212  Last accessed 29.06.2018.  
4 HPQs, inverted exclamatives and rhetorical questions

Cases such as the short attested exchange in (24) might at first sight appear to contradict the claim made in section 2.1 above to the effect that while HPQs can be responded to, they are incompatible with an answer that provides a piece of new information. As A’s utterance in (24) contains the string wel niet, shown above to be incompatible with HDQs (cf. (23b)), this must be a HPQ rather than a HDQ and indeed, under the most natural interpretation of this statement, A wishes to convey the message that it has indeed been a very long time since they played this particular board game. Nevertheless, B’s response appears to constitute an answer which provides a rough indication as to when they last played this game.

(24) A: “Pim-pam-petten? Hoe lang is dat wel niet geleden?”
Pim-pam-pet.INF how long is that PRT not passed
‘Pim-pam-pet [a board game, nn & ld], how long ago is that’
B: “Een jaar of dertien.” zei ik droogjes.
A year or thirteen said I dryly
‘About thirteen years, I said dryly.’

We see two possible ways to account for such cases. The first is to follow Delfitto & Fiorin (2014) who claim that while exclamatives and rhetorical questions have essentially the same (Boolean) semantics (regardless of whether an expletive negator is present), the two structures are differentiated by – among other things – the fact that only the latter can be answered in a pragmatically felicitous manner (see in particular Delfitto & Fiorin 2014: 293, fn. 2). If this is indeed the case, then considering HPQs to be rhetorical questions potentially accounts for the fact that HPQs pattern with exclamatives rather than with HDQs in the ways described above, and yet nevertheless permit a contentful answer like HDQs and other questions, and unlike canonical non-inverted exclamatives.

Note however that – as shown by the constructed dialogue in (25), where A is Jespersen’s example of an English exclamative structure involving expletive negation (discussed above, where it is provided as example (14)) – it is also possible for an interlocutor, B, to provide a pragmatically felicitous answer that provides new information, something which is impossible in response to the corre-

---

192

---

19

https://mijn.editio.nl/schrijfwedstrijd/schrijverij/
sponding non-inverted exclamative in the parallel dialogue in (26).

(25) A: How often have I (not) watched him!
    B: I don’t know. 3 times maybe?

(26) A: How often I have watched him!
    B: #I don’t know. 3 times maybe?

Should the label ‘rhetorical question’ thus be extended to all inverted exclamatives? This seems undesirable, in the first instance as not all rhetorical questions can have exclamative interpretation, and so the affinity which HPQs and other inverted exclamatives show to non-inverted exclamatives is obfuscated under such an approach. On the contrary, canonical rhetorical questions are rather understood as ‘an assertion of opposite polarity’ (Sadock 1971: 224), such that (27) and (28) are interpreted as meaning roughly ‘No-one still uses Facebook’ and ‘Everyone likes chocolate’ respectively.

(27) Who still uses Facebook?

(28) Who doesn’t like chocolate?

Similar examples of canonical rhetorical questions which systematically resist exclamative interpretation can be found in Standard Dutch (where such rhetorical questions can be differentiated from questions with interrogative illocutionary force by the addition of the unstressed discourse particle nou (in Northern Dutch varieties), or nu (which is more idiomatic in Southern Dutch)). A relevant example is given in (29) (from Dik 1997: 244, his (30b) (our translation rm & ld, emphasis in original)), which again has opposite polarity interpretation (‘No-one would want to be chairperson’):

(29) Wie wil er nou voorzitter worden???
    ‘Who on earth would want to be CHAIRPERSON?’

As Nye (2009: 26 fn. 11) observes for English, the same applies to certain structures involving how, such as (30), which are interpreted as meaning ‘It can’t be so very difficult!’ and can thus also be classified as rhetorical questions. Similarly,

---

20 As Delfitto & Fiorin (2014: 296) observe, “Typically, negative rhetorical wh-questions express some sort of universal quantification.”

21 As suggested by the capitalisation, (29) is most naturally pronounced with heavier-than-neutral stress on the predicate voorzitter ‘chairperson’.

193
the Dutch rhetorical question (featuring the particle *nu*) in (31) has very much the same meaning as its English counterpart in (30).

(30) How difficult can it be?
(31) Hoe moeilijk kan dat *nu* zijn, cijfers met elkaar verbinden
how difficult can that PRT be.INF numbers with each.other connect.INF
tot een tekening. *22*
to a drawing
‘How difficult can that be, connecting numbers to make a drawing?’

As neither HPOs nor inverted exclamatives of the type illustrated in (25) are interpreted as assertions of the opposite polarity, there seems little motivation on interpretive grounds for categorising these together with core rhetorical question cases such as (27)–(31). Similarly, in the absence of any clear definition which identifies and characterises properties common to all of the many structurally and interpretively heterogeneous structures to which the term ‘rhetorical question’ has been applied (for various approaches, see e.g. Han 2002, Caponigro & Sprouse 2007, Sprouse 2007, Delfitto & Fiorin 2014, and references cited therein), extending the reach of this label to cover HPOs (and potentially other inverted exclamatives too) currently offers little by way of explanatory advantage. *23*

A starting point for a second – and in our view more convincing – approach to accounting for cases such as (24) and (25) is the fact that what these examples hold in common is that in both cases B wilfully misconstrues what A intends to be an inverted exclamative as a degree question. B’s characterisation of their own manner of responding in (24) as *droogjes* ‘dryly’ is telling in this regard, and B’s response in (25) comes across as equally flippant. We suggest that in (24) and (25) the interlocutor B is in effect providing an answer to the HDQs ‘How long ago is that?’ and ‘How often have I watched him?’ respectively, rather than an appropriate response to the HPQ ‘How long ago is that!’ in (24) and the inverted exclamative ‘How often have I watched him!’ in (25). Thus we can maintain the claim made in section 2.1 that HPOs, like exclamatives, do not introduce a question into the discourse and are thus incompatible with an answer that provides a new piece of information. Speakers can, however, knowingly play upon the surface similarities HPOs show to HDQs and choose to treat HPOs as HDQs, to


*23* Delfitto & Fiorin (2014: 293, fn. 2) themselves “leave the distinction between exclamatives and rhetorical questions to the reader’s intuition”.

194
Thus, whilst recognising the properties which (certain types) of rhetorical questions seem to share with (certain types of) exclamative structures (as identified by Delfitto & Fiorin 2014), pending a clear definition of the term ‘rhetorical question’ and a thorough exploration of the intersection of this category with that of (inverted) exclamatives, we continue to treat HPQs as a particular subtype of inverted exclamative, on the basis of their syntactic and interpretive similarities to canonical members of the category ‘exclamative’, illustrated and discussed in sections 2 and 3 above.

5 For future research

Many questions remain concerning the specifics of the structure of HPQs, in particular how to encode structurally not only the syntactic and semantic differences they show to the sometimes string-identical HDQs, but also the more subtle interpretive differences they show to canonical exclamatives, to other inverted exclamatives, and to certain types of rhetorical questions. As discussed in section 4, a starting point for this would be a detailed taxonomy of the various structures labelled as ‘rhetorical questions’ and ‘inverted exclamatives’ in the literature, with particular attention paid to any potential intersection between these two categories.

It also remains to be determined what the precise function and interpretation of (wel) niet is in Dutch HPQs, in comparison to those Dutch HPQs in which expletive negation is not present. A broader question in relation to expletive negation concerns the range of structural environments in which this is licensed. It is currently unclear why expletive negation in Dutch is permitted in HPQs as well as other canonical exclamatives, while in English it is excluded from HPQs and the majority of other exclamative structures, but permitted in some other con-
texts, including at least marginally in exclamatives such as (14) (repeated above as (25A)).

At least in Dutch, it does seem to be the case that expletive negation in exclamatives is parasitic on the presence of one crucial ingredient, namely wh-movement to the clausal left periphery. In wh-exclamatives lacking wh-movement, expletive negation is totally unacceptable. One such type of exclamative is found in the native variety of Dutch of one of the authors of this squib, which is a variety of East Flemish from the region of Ghent. The relevant structure features an invariable (and typically phonologically reduced) wh-word wa(t) ‘what, which is not located in the left periphery of the clause, but rather sits in the TP. Consider the examples in (32):

(32) a. Gent is wa de max
   Ghent is WHAT the max
   ‘Ghent is SO cool!’

   b. Oostende is wa de max
   Ostend is WHAT the max
   ‘Ostend is SO cool!’

   c. dat is wa schoon
   that is WHAT beautiful
   ‘That’s SO beautiful!’

   d. Wajoooooo da is wa lang geleden
   EXCLAM that is WHAT long passed
   ‘Woah, that’s SUCH a long time ago!’

The structure illustrated in (i) (see Huddleston & Pullum 2002: 845–846) is another context where expletive negation is possible in English. Both interpretations (a) and (b) are available for the string in (i): reading (a) results from interpreting (did)n’t as having negative force, while reading (b) results when (did)n’t is interpreted as expletive negation.

(i) I wouldn’t be surprised if they didn’t get the job.
   a. I would not find it surprising if they did not get the job.
   b. I would not find it surprising if they did get the job.
Crucially, although the negator *nie(t)* can be added to these structures, it always has to be interpreted as inducing canonical sentential negation:

\[(33)\]

a. Gent is **wa** nie de max!
   Ghent is WHAT not the max
   ‘Ghent is SO not cool!’

b. Oostende is **wa** nie de max!
   Ostend is WHAT not the max
   ‘Ostend is SO not cool!’

c. dat is **wa** nie schoon!
   that is WHAT not beautiful
   ‘That’s SO not beautiful!’

d. Wajoooooo dat is **wa** nie lang geleden!
   EXCLAM that is WHAT long passed
   ‘Woah, that’s SO not a long time ago!’

To conclude this section, while *wh*-movement may be a necessary condition for the licensing of expletive negation, the presence of *wh*-movement alone cannot be a sufficient condition for this, given that canonical interrogatives – both in Standard Dutch and in the colloquial variety from Ghent in which structures like [32] are productive – also involve *wh*-movement and yet fail to allow expletive negation, as example [23] above shows.

Providing an answer to the questions raised in this section goes well beyond the scope of our contribution here, which hopes to serve as the impetus for further research in these areas.

### 6 Conclusion

In the course of this paper, we have not only demonstrated that Dutch HPQs pattern alike with exclamatives rather than HDQs on the tests proposed by [Nye27](https://twitter.com/genteswademax). Last accessed 09.07.2018.


[Nye29](https://www.facebook.com/JPBauwens/posts/1052229218190696?comment_id=1052428764837408&comment_tracking=%7B%22tn%22%3A%22%22%22%22%22%22%22%). Last accessed 01.07.2018.


197
for English but, more significantly, have shown that the Dutch data reveals an additional similarity between HPQs and exclamatives to the exclusion of HDQs, namely the ability for expletive negation to occur in the former two structures, but not the latter. This provides additional support for an analysis whereby HPQs differ structurally from HDQs, as opposed to the alternative in which the interpretation of a single interrogative structure is determined by pragmatic context. *Hoe cool is dat wel niet!*

References


Polarity and other distributional properties of uitkijken/opletten/oppassen (‘look out’)

Albert Oosterhof

1 Introduction

This short article presents and discusses some results of a corpus investigation into the use and distribution of the verbs uitkijken, opletten and oppassen (‘watch out, look out, take care’). The study focuses on some aspects of the distribution of these verbs which are relevant to the study of negative polarity items. I will show that the verbs under consideration can be described as negative polarity items which are licensed or triggered by conditions that are applied cumulatively.

Hoeksema (1999) argues that for some negative polarity items, the conditions which “license” or “trigger” polarity items are applied cumulatively. He discusses the properties of the verb wijsmaken (‘deceive’) and argues that this verb shows a special type of sensitivity to negation. Hoeksema makes a distinction between two uses of this verb wijsmaken. First, there is a so-called neutral use, as illustrated in (1). In this use, the verb refers to a situation in which someone really deceives someone into thinking something. We can add a negative adverb to such sentences, as is illustrated in (1b) and the two sentences in (1) are both acceptable.

(1) a. We hebben hem wijsgemaakt dat het feest is uitgesteld.
   ‘We have deceived him into thinking that the party has been postponed.’

An earlier version of this material was presented on July 9 2018 as part of an informal colloquium series during the academic year 2008-2009, organized by Liliane Haegeman at Ghent University. Some of the results were published in an article written in Dutch (Oosterhof & Hoeksema 2008). Author’s address: KU Leuven, Sint-Andriesstraat 2, 2000 Antwerp, albert.oosterhof@kuleuven.be.
b. We hebben hem (toch maar) niet wijsgemaakt dat het feest is uitgesteld.
   ‘We have not deceived him into thinking that the party has been postponed.’

Secondly, the verb can be used in counterfactual environments as illustrated in (2). The relevant reading is available in sentences containing a negative element, like (2b). Hoeksema (1999) observes that this counterfactual use does not seem to be available if the adverb of negation is absent, as in (2a).

(2) a. #Je maakt mij niet wijs dat het feest is uitgesteld.
   (Suppose someone tried to deceive me into thinking that the party has been postponed, then I would not believe it. \( \approx \) ‘I do not believe that the party is postponed.’)

b. Je maakt mij niet wijs dat het feest is uitgesteld.
   (Suppose someone tried to deceive me into thinking that the party has been postponed, then I would not believe it. \( \approx \) ‘I do not believe that the party is postponed.’)

Hoeksema (1999) concludes from such observations that *wijsmaken* can be described as a negative polarity item with triggering conditions that can be applied cumulatively. This item is not a “standard polarity item”, but in a certain environment, in this case in counterfactual contexts, it requires the presence of negation.

My corpus research presented here focuses on the distribution of the verbs *uitkijken*, *oppassen* and *opletten* and on similarities with the distribution of *wijsmaken*. The sentences presented in (3) are corpus examples taken from the Spoken Dutch Corpus, a corpus of some ten million words with spoken Dutch material.

(3) a. en als je dan niet uitkijkt dan uh dan gaat ‘t één ten koste van ‘t ander. (CGN)
   ‘if you do not look out, then one thing is at the expense of another thing.’

b. nou je kan er ruzie om krijgen als je niet oppast. (CGN)
   ‘in fact there will be a row if you do not look out.’

c. en dan uh uh mmm dan wordt ‘t uh al gauw een puinhoop als je niet oplet. (CGN)
   ‘it all becomes a mess quickly if you do not look out.’
In these sentences, the verbs under consideration are used in a conditional clause and accompanied by an adverb of negation. Intuitively, such conditional sentences with verbs like *uitkijken* are less natural without the adverb of negation. The pattern in (4a), in which an adverb of negation is used, is a more natural type of conditional sentence than (4b).

(4) Intuition:
   a. Als je *niet* uitkijkt/oppast/plet, dan X
      'If you do not look out, then X' (= a very 'natural' type of sentence)
   b. Als je uitkijkt/oppast/plet, dan X
      'If you look out, then X' (= a less 'natural' type of sentence)

I do not claim that there is a strict difference in grammaticality or acceptability between the patterns in (4). I will carry out a corpus investigation to verify the introspection-based pattern in (4). This method is in line with current research into negative polarity items, in which claims about the polarity sensitivity of items are based on such corpus data.

2 Corpus

Table 1 presents an overview of corpus material used in my corpus study. The material consists of various types of texts, spoken as well as written material; formal as well as informal texts and Flemish as well as Dutch texts. The relevant information can be found in the table.

3 Results

The results of the corpus study are presented in Table 2. Only occurrences where the verb has the desired meaning are taken into account. For example, *uitkijken* has an alternative meaning under which the verb can be translated in English as 'look forward to'. We are only interested in cases where the verbs under consideration can be translated as 'watch out', 'look out' or a similar expression.

The three verbs under consideration have a relatively similar distribution. For each verb the occurrences in modal contexts and imperatives correspond to two thirds or more of the total number of sentences.

Each verb occurs in combination with a 'nonveridical' (cf. Zwets 1995) trigger in more than 90% of the total number of cases. In the literature, such items
Corpus Number of Period Material
Words


CGN (2004) 10 million 1998-2004 Dutch and Flemish spoken material: dialogues, other conversations, interviews, lessons, speeches, news bulletins, live coverages and (other) spoken texts

CONDIV-corpus 4.8 million 1998 material from Dutch newspapers

Mediargus (online digital archive of articles from Flemish newspapers and magazines) relevant number of words cannot be determined material from Flemish newspapers and magazines

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Number of Words</th>
<th>Period</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGN (2004)</td>
<td>10 million</td>
<td>1998-2004</td>
<td>Dutch and Flemish spoken material: dialogues, other conversations, interviews, lessons, speeches, news bulletins, live coverages and (other) spoken texts</td>
</tr>
<tr>
<td>CONDIV-corpus</td>
<td>4.8 million</td>
<td>1998</td>
<td>material from Dutch newspapers</td>
</tr>
<tr>
<td>Mediargus (online digital archive of articles from Flemish newspapers and magazines)</td>
<td>relevant number of words cannot be determined</td>
<td>1990-2008</td>
<td>material from Flemish newspapers and magazines</td>
</tr>
</tbody>
</table>

Table 1: Overview of the corpus material

which show a (strong) tendency to occur in nonveridical environments are referred to semi-NPI’s or weak NPI’s (see Hoeksema 1994, van der Wal 1996, Oosterhof 2003-4). On the basis of our results, we could draw the conclusion that uitkijken, oppassen and opletten are polarity sensitive items on the basis of their general distribution. The question is of course whether this is the desired result, but I will return to this question in the discussion section.

Now that we have discussed the general distribution of the three verbs, we will now consider some more detailed results with regard to conditional sentences. Table 3 shows a comparison between the use of the verbs in conditional sentences and other contexts. The table provides for each case the number of occurrences in combination with an adverb of negation niet and the number of occurrences without such an adverb of negation.
### Table 2: Distribution of the three verbs

<table>
<thead>
<tr>
<th>environments/triggers</th>
<th>uitkijken</th>
<th>oppassen</th>
<th>opletten</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>look/watch out, take care</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>moet (‘must, have to’)</td>
<td>137</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>mogen (‘may, be allowed to’)</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>kunnen (‘can’)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>zullen (‘shall’)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>modal contexts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>modal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>auxiliary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imperatives</td>
<td>niet (‘not’)</td>
<td>55</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>niet goed (‘not well’)</td>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>negative contexts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conditional sentences</td>
<td>als…niet…(‘if…not…’)</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>als…niet goed (‘if…not well…’)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>wennear + niet (‘when + not’)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>mits…(‘only if…’)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>inherently negative</td>
<td>zonder (‘without’)</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>negative</td>
<td>alleen (‘only’)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>future</td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>questions</td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>combination of two ‘nonveridical’ triggers</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>other ‘nonveridical’ triggers</td>
<td>4</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>without any trigger</td>
<td>14</td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td>TOTAL</td>
<td>281</td>
<td>384</td>
<td>419</td>
</tr>
</tbody>
</table>

There is a significant difference between the use of these three verbs in conditional sentences and other contexts. In conditional sentences, by far most occurrences of the verbs are accompanied by the adverb of negation. In other con-
texts, most sentences in which one of the verbs is used do not contain an adverb of negation

4 Discussion

More than 90% of the occurrences of the verbs uitkijken, oppassen and opleten can be identified as occurrences in nonveridical contexts, which implies that these verbs can be described as weak negative polarity items. However, this conclusion is in a sense counterintuitive: there is only a small proportion of the total number of occurrences in which an adverb of negation is used (cf. Table 2). The nonveridicality hypothesis leads to a situation where an item can be identified as a (weak) negative polarity item, even though the percentage of occurrences with negation is low. If we accept the nonveridicality hypothesis, the relation between the set of items identified by using the relevant criterion and the concept of polarity sensitivity is lost. The criterion of nonveridicality provides us with a very broad definition, as a consequence of which even items not especially sensitive to polarity end up as polarity sensitive items.

In earlier work on negative polarity items other logical properties were used in order to describe the distribution of negative polarity items. In publications from the eighties and the nineties by people like Ladusaw, Zwarts, Hoeksema and van der Wouden the assumption was made that negative polarity items are sensitive to downward entailment and related properties. Modal contexts and imperatives do not have the property of downward entailment. Under such a hypothesis, uitkijken, oppassen and opleten would not be labelled as negative polarity items, which seems to be a more desired result.

Let us now consider the relation between the use of uitkijken, oppassen and opleten in conditional sentences and the presence of an adverb of negation. We have observed that there is a clear tendency that in most cases such conditional sentences contain an adverb of negation. These verbs are polarity sensitive when used in conditional sentences. This leads to a confusing situation. The claim that the verbs uitkijken, oppassen and opleten are negative polarity items in conditional sentences suggests that they must be used with an appropriate trigger. However, the protasis of a conditional sentence is assumed to be an appropriate trigger for negative polarity items itself. This implies that the claim that uitkijken, oppassen and opleten are negative polarity items in conditional sentences is tautologically true (i.e. it cannot be falsified).

This illustrates that there is a complex relation between the sensitivity of so-
called polarity items and their licensing environments. This point can be further demonstrated by referring to another item, *iemand gegeven zijn* (‘be granted to someone’). This item is included by Hoeksema (2013) in his lexicon of polarity items. Hoeksema (2013: 42) writes that this item is fairly common in conditional sentences (5% of total number of attestations). We can conclude that the protasis of a conditional sentence is an example of an environment which plays a different role with regard to different polarity items. It is part of the set of licensing conditions of an item like *iemand gegeven zijn* (‘be granted to someone’), but for *uitkijken*, *oppassen* and *opletten*, it is part of the definition of the relevant item.

As a conclusion, note that we can describe the restrictions on the verbs *uitkijken*, *oppassen* and *opletten* as a special kind of collocational behaviour. A collocation can be defined as an idiosyncratic restriction on the compatibility of lexical items. In our data, there seems to be collocational behaviour as well, but in this case there is an idiosyncratic restriction on the compatibility of environments in which items occur. The three verbs under consideration in our study provide an example of the fact that description of concrete behaviour of lexical items is a more complicated task than expected from the perspective of straightforward logical notions like nonveridicality.

**References**


The absence of classifiers in numeral classifier constructions in Vietnamese

Trang Phan*

To Liliane, who inspires me to always look beyond assumptions

Vietnamese is a so-called obligatory numeral classifier language. Nevertheless it is possible to have direct combination between a numeral and a bare noun in the absence of a go-between classifier. The goal of this squib is to investigate the seemingly unusual cases of such Numeral-N phrases.

In [GIl's (2008)] typology of numeral classifiers, the world's languages are divided into three types on the basis of whether classifiers are absent, optional or obligatory in numeral constructions. Vietnamese is considered as belonging to the third type in which a numeral cannot quantify a noun without the presence of a classifier based on examples like (1):

(1) hai *(con) chó
two CLF dog
‘two dogs’

(Gil's example 2008: 4)

Gil notes, though, that in Vietnamese there is a specialised style of speech, namely food ordering at food stalls and restaurants, in which numeral classifiers are frequently omitted, as seen in (2).

(2) Context: At the noodle bar where there is a choice between two kinds of noodle soup (chicken or beef), a group of 5 customers might simply order as follows:

Ba gà hai bò
Three chicken two beef
‘Three bowls of chicken soup and two bowls of beef soup.’

*My special thanks go to Eric Lander for proofreading and commenting on the earlier version of the squib. Needless to say, all errors are mine.
Gil excludes those specific contexts and still classifies Vietnamese as an obligatory numeral classifier language (Gil 2008: 5). In this squib, I take issue with Gil’s classification by showing that the optionality of Vietnamese classifiers in numeral constructions is not limited to those special contexts, rather the optionality of classifiers in Vietnamese numeral constructions can be understood in a systematic and productive way. More precisely, there is in fact a large group of Vietnamese nouns which do not need a classifier in combination with a numeral.\footnote{See Gil (2008), Aikhenvald (2000) and Greenberg (1974) for similar phenomena cross-linguistically.}

In the literature on Vietnamese, it has been proposed that Vietnamese nouns can be divided into two main types: classified nouns that require a classifier to be quantified (as in (3)), and non-classified nouns that can be counted directly in the absence of classifiers (Emeneau 1951, Thompson 1965, Nguyễn 1975, Nguyễn 2002, Tran 2011, a.o.), as illustrated by the Numeral-N phrases in (3).

\begin{enumerate}
\item \textit{Hai ngày/tỉnh/ bàn/ túi/ bạn/ sinh-viên/vận-động-viên two day/ province/ table/ bag/ friend/ student/athlete ‘two days/provinces/tables/bags/friends/students/athletes.’}
\item \textit{Một-trăm ngày/tỉnh/ bàn/ túi/ bạn/ sinh-viên/vận-động-viên one-hundred day/ province/ table/ bag/ friend/ student/athlete ‘one-hundred days/provinces/tables/bags/friends/students/athletes.’}
\end{enumerate}

This phenomenon is not novel cross-linguistically. The same Numeral-N pattern is found in Korean (Hwang 2012) but with two restrictions: Numeral-N is only possible if the NP (i) refers to common human nouns and (ii) is combined with numerals below five, as seen in (4).

\begin{enumerate}
\item \textit{*tases kapang five bag ‘Five bags.’}
\end{enumerate}

\footnote{Her et al. (2015) also observes that Numeral-N phrases can be licensed in Chinese in limited contexts, in an idiom for instance:}

(i) \textit{wu ma huan liu yang 5 horse trade 6 goat ‘Trading 5 horses for 6 goats.’}
Unlike in Korean, Numeral-N phrases in Vietnamese are productive regardless of the noun type and the numeral type, as shown in (3), which clearly suggests that, contra Gil’s classification, Vietnamese seems to better fit with optional numeral classifier languages.

In fact, a Numeral-N phrase can appear as the object or the subject of sentences, and can be interpreted as indefinite or definite in Vietnamese:

(5) Tôi vừa viết thư giới thiệu cho hai sinh viên. Hai sinh viên đều học ngôn ngữ học.

‘I have just written recommendation letters for two students. The two students both study linguistics.’

That is to say, a Numeral-N phrase can have the same distribution and interpretation as a fully-fledged Numeral-CLF-N phrase:

(6) Tôi vừa viết thư giới thiệu cho hai bạn sinh viên. Hai bạn sinh viên đều học ngôn ngữ học.

‘I have just written recommendation letters for two students. The two students both study linguistics.’

This seems to suggest that a Numeral-N is indeed a Numeral-CLF-N underlyingly. The question is where the covert CLF component resides: in the Numeral or in the N?

One account is proposed by Nguyễn (2002) in which the numeral in Numeral-N phrases is considered as a ‘zero classifier’ (in his terminology) carrying the individuating function. Consider Nguyễn’s minimal pair in (7):

3 In order to account for the optionality of the classifier in numeral constructions, Borer (2005)
In the absence of the numeral, the bare noun in (7a) is interpreted as generic, whereas the addition of the numeral hai ‘two’ in (7b) forces a specific reading. Putting it differently, the absence of a classifier in Numeral-N phrases, according to Nguyễn (2002), results from the fact that the numeral serves two functions at the same time: to individuate and to count.

A scrutiny of the Vietnamese data, however, reveals that this analysis cannot be true since apart from numerals, the bare noun can also co-occur with a number of other elements in the absence of classifiers: it can immediately follow a quantifier (8a) or a plural marker (8b), and can directly precede a demonstrative (8c) or a relative clause (8d).

That is to say, the absence of the classifier in the Vietnamese Numeral-N construction is derived from the fact that the bare noun is underlyingly a CLF-N phrase (along the lines of Cheng & Sybesma’s 1999 proposal for Chinese). That is, rather than assuming that numerals, quantifiers, plural markers, demonstratives and the RC marker all individually combine with a zero classifier, we can take the much simpler approach that it is just the noun that combines with the

117–118) also suggests that in such languages, numerals can function as dividers.
individuating classifier.

This account is further supported by the fact that the bare noun indeed can function as either the object or the subject of the sentence, and can obtain either indefinite (underlying bare N) or definite (underlying CLF-N) readings, as in (9):

(9) Tôi mời cả sinh viên và giáo viên đến dự tiệc. sinh viên thì творить all student and teacher come attend party Student TOP say bị tì, giáo viên thì còn tỉnh táo. drunk TOP teacher TOP still awake

‘I invited both students and teachers to come to the party. The students were very drunk, the teachers were still awake.’

To conclude, in this squib, I have contested Gil’s classification of Vietnamese as an obligatory numeral classifier language based on a wider range of empirical patterns in which the overt classifier is absent. I further put forward the idea that the absence of the classifier in Numeral-N constructions is indicative of the fact that the bare noun is not so bare in Vietnamese.

References


Nguyễn, T.C. 1975. *Từ loại danh từ trong tiếng hiện đại* [The word class of nouns in modern Vietnamese]. Hà Nội: Nhà Xuất bản Khoa học xã hội.


Embedded interrogatives as free relatives

Cecilia Poletto and Emanuela Sanfelici

1 Introduction

In this work we intend to show that a typical property of North-Eastern Italian di-alects (NEIDs) namely the sequence wh-complementizer in embedded interrog-atives (which in some dialects also extends to non-standard (under the definition by Obenauer 2006) and even standard main questions originally developed out of a free relative structure where the complementizer is actually not a comple-mentizer but part of the internal structure of the wh-item heading the relative clause. The structures we are going to investigate are illustrated in (1), where we see that the wh-item is followed by the item che, which has always been inter-preted as a complementizer.

(1) No so cossa che el gabia dito.
   Not know what that he has said
   ‘I do not know what he said.’

If we consider this phenomenon in traditional terms, it looks like an apparent violation of the traditional “doubly filled comp filter”, since the C⁰ head is occu-pied by the complementizer and specifier of the CP by the wh-item. Even un-der a more restricted view, namely that the doubly filled comp filter can only be violated if the head and the specifier agree (see Koopman 1996), there is no morphological hint of an agree operation between the wh and che that might justify the violation. This structure is problematic even in a more modern frame-work: for instance the exceptionality of these structures has been noted (see Poletto & Vanelli 1999 and Poletto 2000) within the cartographic framework as cases in which the complementizer does not sit in its usual Force⁰ position as it does in declaratives (see Rizzi 1997), but as a case of a low complementizer probably located in the low Fin/Focus area. Even assuming that the comple-mentizer is first merged in the Fin⁰, since it has [+finiteness] features and then
raises to Force⁰, we have a case in which the complementizer does not reach the Force⁰ position. Alternatively, we have to assume that the Force⁰ of interrogatives is lower than the Force⁰ of declaratives, which conceptually empties the notion of Force⁰ as marking sentence type. Furthermore, the possibility to have at the same time high and low complementizers occurring after left dislocated items and wh-items has been identified as a weakness of the cartographic approach (see van ?), where ideally every element should have its own position. Furthermore, che is usually the complementizer of declarative clauses, while interrogative clauses (at least yes/no interrogatives use se 'if'). Here we intend to explore an alternative which only recently became available since Poletto & Sanfelici (2015) have argued that the “complementizer” che found in relative clauses in Italian is actually a determiner-like portion of the internal head of the relative clause. We will argue that the first instances of wh-che structures in embedded interrogatives are actually free (or light headed) relative clauses, so that the element che found in these cases is not a complementizer at all, but is part of the internal structure of the wh-item. In order to show how this might work, we will take various NEIDs into consideration.

2 Venetian

In Venetian, and in the majority of Veneto dialects, the sequence [wh-che] is obligatory with all wh-items in embedded interrogatives with no exception. The lack of che results into ungrammaticality. We provide here some examples to illustrate the point. These structures are so stable that they can occasionally leak into the regional standard Italian of native Veneto speakers:

(2) No so dove *(che) el sia 'nda.
   Not know where that he is gone

(3) No so chi *(che) le gabia visto.
   Not know who that them have seen

(4) No so come *(che) el lo gabia fato.
   Not know how that he it have done

(5) No so parcossa *(che) el sia ndá via.
   Not know why that he be gone away

(6) No so quanti pomi *(che) el gabia comprá.
   Not know how many apples that he have bought
Interestingly, Old Venetian does not possess this property in the same way the modern variant does. The following examples are all extracted out of the Tristano Veneto, a medieval text dating around the end of the XIII or the beginning of the XIV century. They show that in Old Venetian it is indeed possible not to have che after then wh-item in embedded interrogatives:

(8) cercando qui novelle li savessè dir del chavalier che
looking for which news them could say of the knight that
persegiva la Bestia Gladisschante
chases the Bestia Gladisschante (370, 32)

(9) domandé in que logo die eser questa bataia.
asked in which place should be this battle (69, 45)

(10) Ma atanto me dixé in qual parte se va li cavalieri.
But in as much to me said in which PART CL go the knights (145, 23)

(11) et se algun me domandasse de qui era la nave
and if some to me ask of whom was the ship (168, 8)

This is generally true of complex wh-phrases but also of bare wh-items:

(12) Lo re domandà Tristan qui era lo chavalier che chavalchava
The king asked Tristan who was the knight that was riding
si solo
alone (89, 47)

(13) Et ello li domandá donde elli era
And he to them asked where he was (165, 13)

(14) io non so qui elli sia questi tre cavallieri de que le lettre
I not know who they are these three knights about whom the letter
parla
talks (22, 14)

Notably, the element meaning where, besides the form in [13] also has a different structure, which looks like a free relative clause with a light head, like the pronoun ‘there’ lá. The same is true of the item ‘why’, which displays the structure of a relative with the light noun ‘reason’ chasion:
and asked there where that he was the king (42, 3)

She to them asks the reason why they were come (407)

The same can also occur with the wh-item *qui*, ‘who’, which has no pronoun or light noun, but can (though it need not, cf. (12) be followed by *que*:

wanted to know who that he was

Notice that these cases have to be analyzed as embedded interrogatives, since they are selected by the verb *domandar* ‘ask’ and *saver* ‘to know’. However, the case in (18) shows that the light head represented by the deictic pronoun *lá* is not necessary to have the structure *o-que*. This pattern resembles that exhibited by free vs. light-headed relative clauses: a free relative clause can have a null head as in (18) or a light-headed one as in (15) (Poletto & Sanfelici 2015, 2018).

The phenomenon *wh-če* in Old Venetian is thus found primarily with ‘where’, ‘who’, ‘what’ and ‘why’, and in this last case only with a light headed relative clause, while the other wh items can also have a free relative tout court. The sequence is not found with ‘how’, and ‘from where’ or with complex wh phrases in the Tristano Veneto. We can conclude that Old Venetian displays the first stage of the phenomenon which then extends to all wh-items, i.e. embedded interrogatives actually display the structure of a free relative clause with a light head noun and the following *če*.

3 Trentino

One further dialectal area that presents this phenomenon, although in a more limited way, is Trentino: here not all wh-items have to be followed by *če*, but only some of them. Garzonio (2007) already notices this fact on the basis of the survey of the ASIt database:
The realization of *che* is indeed variable, since the same speaker translates sentences with and without *che* for the same wh-item in the prompt.

The same observation is proved by a further investigation based on her own field work by Polonia (2014) for the Trentino dialects spoken in Val di Sole, Val di Cembra and Val di Non. She notices that the occurrence of wh-*che* is not obligatory with various wh items, but there is a clear tendency to use *che* with the following wh-items:

(20) Dime kel *che* le te a dit le to sorele.  
Tell-me that that cl to. you have said your sisters [Tassullo]

(21) No so chi *che* lava giò i piatti.  
Not know who that washes down the dishes [Stedron, Segonzano]

Notice that the structure with the wh-item corresponding to ‘what’ in (20) is the same used in French embedded interrogatives where there is no wh-items but a light headed relative clause whose head is still clearly the distal demonstrative (*ce que*). Polonia (2014) provides a hierarchy of wh-items which can more frequently be doubled by *che*, although she also notes that the same speaker varies in using *che* after the very same wh-item or not. A simplified summary of the implicational scales she reports for the three valleys she investigates is reported here:

(22) What/where who how many which when which X how many X why/how

The data from Old Venetian and those from Trentino only partially overlap. We can recognize some tendencies, since in all the varieties the elements which display the doubling are where, who, what but Old Venetian also tolerates why (with the light noun chasion) while complex wh-phrases are clearly at the right edge of the spectrum and the element corresponding to ‘how’ is also generally
4 Preliminary notes for a syntactic analysis

Poletto & Vanelli (1995) propose the following generalization:

(23) If a dialect displays wh-*che* in main interrogatives it does also in embedded interrogatives.

Here we propose that this generalization can also be read in terms of diachronic development, such that wh-*che* starts out in embedded interrogatives and then expands into main interrogatives. Poletto & Vanelli single out the last step of the development of this structure, namely the point when it extends from embedded to main interrogatives, which we will not deal with here (see Poletto 2000 for an analysis in cartographic terms). At this point two questions arise: (i) how do these structures start out in embedded interrogatives? (ii) When we take into consideration the data presented in the previous section where we showed that the presence of the wh-*che* in embedded interrogatives depends on the type of wh items introducing them, how can the generalization in (23) be implemented.

Addressing the first question, we propose that the trigger for the rise of wh-*che* is the semantic ambiguity which can arise between free and interrogative clauses under certain predicates (see Caponigro 2003). Parry (2003) shows that in Old Piedmontese the first attested cases of wh-*che* are precisely ambiguous between a free relative and an interrogative reading. One might thus simply say that there was a semantic ambiguity between a free relative and an embedded interrogative (along the lines of Cecchetto & Donati 2015) and this is the reason why the structure wh-*che* could be extended from free relative clauses to embedded interrogatives. However, we think that the question is more complex than that, and that the first link towards the intrusion of a relative structure in an interrogative one is not between embedded interrogatives and free relatives altogether but between embedded interrogatives and light headed free relatives. Free relatives can either be expressed by a simple wh item, as shown in (24), but Italian varieties (as other Romance languages) have the tendency to realize light heads, especially with the element corresponding to *what*, which, just like in French, is utterly impossible as a free relative (25) and requires a light head (26) (see Munaro 2000).
(24) Chi ha detto questo, ha sbagliato.
   ‘Who said this was wrong’
(25) *Cosa hai detto, è sbagliato.
   ‘What you said is wrong’
(26) Quello che hai detto, è sbagliato.
   ‘The thing that you said is wrong’

As predicted under our proposal, if wh- *che* embedded interrogatives are free/ light-headed relative clauses, we expect the embedded interrogative with the element corresponding to ‘what’ to be impossible at least in certain varieties. This prediction seems to be born out: in certain Trentino dialects the embedded interrogatives on ‘what’ cannot have the wh-item for ‘what’, but require the light-head *quello che* “that that” as they do in real free relatives (Garzonio 2007). Poletto & Sanfelici (to appear) attribute the spell out of a light head to the fact that, as assumed by Cinque (2013) and originally proposed by Sauerland (1999, 2003), all relative clauses are double headed: they all have an external head located in the DP spine and an internal one located inside the relative clause. Languages differ with respect to the head they lexically spell out, and free relatives can be analyzed as restrictive relatives where neither head is lexically realized. Notice however that there is an “intermediate stage” between real free relatives and headed restrictive relative clauses, namely light headed relatives, which are exactly the type of relatives that first manifest themselves in embedded interrogative contexts. We surmise that the variation found in Old Venetian attests precisely this: the first relative structures to be found in interrogative domains are light headed relatives as the ones found with *là o que* and *la chasion que*. The fact that the first step of the evolution of the wh-*che* structure is a light headed relative clause is shown by the fact that these are the only possible structures even in languages, like French, where no other wh-item presents this possibility. If French represents the first stage of the development of this structure, NEIDs represent the second one: the usage is extended to wh-pronouns, where there is no lexical light head. If this were the end of the story, we would expect that all wh-items should allow for doubling except for the complex ones already containing a lexical head. However, when we take into consideration the data presented in the previous section where we showed that the presence of the wh-*che* in embedded interrogatives depends on the type of wh items introducing them, this expectation is not fully fulfilled. The stage depicted by Polonia (2014) and Garzonio (2007) for Trentino dialects only partially matches this expectation: we
find that the wh-items that tolerate the wh-*che* structure are a) those that still display the light head (like the case of *kel che* in (20)); b) those that allow for a null free relative head are not all possible wh-items, but there is a second hierarchy even if we exclude light headed relatives. The implicational scale in (23) shows that elements like *why* and *how* are much more resistant, even more than complex wh-items. We are forced to assume that there is one more factor at play here. If we leave complex wh-items aside, the hierarchy in (23) partially looks like the Keenan and Comrie’s case hierarchy, which goes from the less complex to the most complex cases: here evidently what, where and who are less complex than why and how, which are at the rightmost edge of the scale. It is well known that the Keenan and Comrie’s hierarchy is at play in relative clauses in the phenomenon that is known as the case containment condition, i.e. the condition ruling which case is spelled out on the lexical relative head when the case of the main clause and the case of the relative clause do not match. Poletto & Sanfeti (to appear) show that modern Italian generally resolves case mismatches in favor of the external case:

(27) Partiró con chi hai incontrato ieri.
    ‘I will leave with whom you met yesterday’

(28) *A chi hai telefonato é partito.
    ‘To whom you phoned left’

This is the case also for the NEIDs we are considering, as the translation in modern Venetian shows:

(29) Ndaró via co chi che te ga visto jeri.
    ‘I will leave with whom you met yesterday’

(30) *A chi che ti ghe ga telefoná el ze partio.
    ‘To whom you phoned left’

Suppose that embedded interrogatives in these dialects have the structure of free relative clauses (as the light headed cases overtly show). This means that they have two heads, a relative internal one, and an external one, which is the (null) DP that is the complement clause of verbs like ‘ask’. If we apply the double-head idea to embedded interrogatives, we immediately explain the first step of the evolution, i.e. light headed relatives, which spell out both the external light head and the internal wh head: *in la chasion perché*, *la* is the determiner of the whole DP, *chasion* is the external head and *perché* represents the internal head
by means of a relative pronoun. Cases like lá o che display the external head (the deictic pronoun lá) and the internal one, which is made up of two components [o che]. This means that at the successive stage of development, i.e. the one of a real free relative with no lexically spelled out external head, we only have the internal one [o che]. One might wonder why an element like che intrudes into a wh-item and what portion of its internal structure it spells out. The fact that che can be part of a wh-item is actually straightforward if we consider the form corresponding to ‘what’ in standard Italian, namely che cosa. Poletto & Pollock (2009) propose the following internal structure for wh-items:

(31) \[
\begin{array}{c}
\text{DisjP} \\
\text{ExistP} \\
\text{RestrictorP}
\end{array}
\]

Suppose that in Italian cosa represents the restrictor, since it literally means ‘thing’ and che the existential component as in (32):

(32) \[
\begin{array}{c}
\text{DisjP} \\
\text{ExistP che} \\
\text{RestrictorP cosa}
\end{array}
\]

This structure immediately explains the complex nature of wh-pronouns and the ordering of the elements we observe. An independent argument in favor of the idea that wh-pronouns have the complex structure illustrated in (32) is that there are dialects where two components, i.e. the restrictor and the disjunctive as in (33) and (34) can occur separately giving rise to doubling structures (see Obenauer 2006, Munaro & Poletto 2018):

(33) Cossa invit-to chi?!
    COSSA invite-you who(m)
    ‘Who (the hell) are you inviting?!’

(34) Cossa ve-to dove?!
    COSSA go-you where
    ‘Where (the hell) are you going?!’
    [Paduan (Central Veneto)]

Notice that our free relative/interrogative case represents the third logical possibility besides the combinations of existential plus restrictor and restrictor plus disjunctive:

(35) \[
\begin{array}{c}
\text{DisjP} \\
\text{ExistP che} \\
\text{RestrictorP}
\end{array}
\]

Notice that also this combination is found in doubling structures:
On this basis we can conclude that the structures [wh che] in embedded interrogatives is not a case of low complementizers but the spell out of the ExistentialP, one of the internal projections of the wh-item. In this sense interrogative and relative pronouns are the same, i.e. they contain at least three layers, which can be lexically spelled out or not. The reason why the first embedded interrogatives to manifest the spell out of the existential component che are actually ambiguous with free relative clauses has to do with the fact that in general in relative clauses ExistentialP is always spelled out by che in Italian varieties.

5 Concluding remarks

In this work we have entertained the hypothesis that the occurrence of che after wh-items found in NEIDs is not a case of violation of the doubly filled comp filter, where the complementizer che sits in a low C° head. We have proposed that these structures, just like relative clauses, are the spell out of more than one internal projection of the wh-item. This explains why the first attested historically cases are indeed free relative clauses, since relative clauses generally spell out the existential portion of the internal relative head. It can also explain well known cases of wh-doubling in simple wh-questions, which are rather rare even in languages like colloquial German and non-standard Dutch varieties which display a so called scope marker wh- doubling the actual wh-item in long extractions.

References


Liliane's academic career and mine have been intertwined for decades. We both have a common obsession with collecting 'real' data on English usage. I must admit, however, that Liliane’s method of collecting data is far more rigorous than mine: she collects examples of interesting structures in written English while reading the Guardian (etc.) on her exercise bike; I collect data on unusual structures in spoken English while lying prostate on a sofa (sporadically dipping into a pot of Ben and Jerry's Chocolate ice-cream) and listening to live commentary on the latest football or cricket matches on radio or TV. We exchange data regularly, as you'll see from numerous footnotes in her excellent (2012) monograph *Adverbial Clauses, Main Clause Phenomena, and Composition of the Left Periphery* citing weird structures that I'd collected (or dreamed up), and from the frequent mentions of her in my forthcoming CUP monograph on *Colloquial English: Structure and Variation*.

Liliane once told me that her interest in generative syntax had been sparked by reading my 1981 *Transformational Syntax* book as a student, and that she'd always wanted to write a joint article with me. Together with a Spanish colleague (Ángel Jiménez-Fernández from Sevilla), we worked on a paper on extraction out of subjects in English and Spanish, which was eventually published with the title ‘Deconstructing the Subject Condition in terms of cumulative constraint violation’ in *The Linguistic Review* (2014, vol. 31, pp. 73–150). The paper combines Minimalist, Cartographic and processing perspectives, and was well received: Chomsky is reported to have liked it, and (at the time of writing) it is the tenth most frequently downloaded article on LR's website.

---

*The most famous being ... no, not Poirot but rather Kevin de Bruyne. After all, he has his own wallpaper!
Over the years, each of us has provided inspiration for the other. For example, in my 2004 *Minimalist Syntax* book (p. 429), I reported the sentence ‘What is thought has happened to him?’ to have been produced by a reporter on BBC Radio 5. This observation prompted Liliane to take a deeper look at this type of structure (in which a *wh*-subject is raised out of a finite clause), culminating in the publication in 2017 of her article on ‘Syntacticizing blends: the case of English *wh*-raising’, written jointly with Lieven Danckaert.

Conversely, Liliane has made numerous insightful remarks which have guided and inspired my own research. For example, my forthcoming CUP monograph on *Colloquial English* has a lengthy chapter on non-canonical uses of the complementiser *that* in spoken English, and includes an appendix on the use of that in adverbial clauses in present-day English in structures such as the following (1b-h being taken from recordings I have made of live, unscripted radio broadcasts):³

(1) a. They got wet [because *that* it was raining] (Alison Henry, pc)
b. The reason that England won’t win the world cup is [because *that* the younger players coming through are too spoiled] (Andy Goldstein, Talksport Radio)
c. Some people were talking about it as some sort of race riot, [as if *that* the Dutch team was split along racial lines] (Andy Brassell, BBC Radio 5)
d. [Although *that* they won the title], they finished on something of a low (John Cross, Talksport Radio)
e. [Even though *that* we lost], I'd still put that down as one of my favourite games (Brett Lee, BBC Radio 5)
f. And [when *that* we were 71 for none], there was a chance to sort of close the game out (Peter Moores, BBC Radio 5)
g. For many it was inevitable, [once *that* David Hay walked into the room] (Mike Costello, BBC Radio 5)
h. That’s been the dominant philosophy in Brazil, [ever since *that* they lost to Holland] (Tim Vickery, Talksport Radio)

Such structures were found in Chaucer, and thus may well be an archaic feature which has gradually been dying out over the centuries, but which has survived in a minority of speakers from a wide variety of backgrounds: Andy Brassell, John

---

³I’m grateful to Cambridge University Press for allowing me to present this short summary of part of the discussion of adverbial clauses in §3.8 of my forthcoming monograph on *Colloquial English*. 227
Cross, Andy Goldstein, Matt Scott and Tim Vickery are Londoners, Mike Costello was born in Bromley, Peter Moores was born in Macclesfield, Graham Beecroft comes from Merseyside, Brett Lee was born in Wollongong (Australia), and Alison Henry comes from Belfast.

Given that much of Liliane’s (2012) monograph was devoted to adverbial clauses, I discussed the data with her. She drew my attention to Lieven Danckaert’s excellent book on *Latin Embedded Clauses: The Left Periphery* (Amsterdam: Benjamins, 2012), and he kindly sent me a copy of the book, and discussed the data with me. Danckaert argues that in adverbial clauses, there is parametric variation in respect of the position of subordinating conjunctions, with some occupying a high position in the periphery, and others a low position. This raises the possibility that the subordinating conjunction/SUB in the adverbial clause in a sentence like (1a) above could occupy a high position in the periphery in minority varieties which allow the use of *that* in adverbial clauses, along the lines shown below:

(2) They got wet [SUBP [SUB because] [FORCEP [FORCE that] [FINP [FIN ø ] it was raining]]]

By contrast, in majority varieties which don’t allow that in adverbial clauses, the conjunction may occupy a lower position in the periphery — perhaps FIN, as below:

(3) They got wet [SUBP [SUB ø ] [FORCEP [FORCE ø ] [FINP [FIN because] it was raining]]]

This would mean that there is no peripheral head position after *because* for *that* to occupy in (3), so accounting for why mainstream varieties don’t allow *that* after conjunctions.

However, there is evidence against analysing conjunctions like *because* as FIN heads. For one thing, a key defining property of low subordinating conjunctions identified by Danckaert (2012) is that they allow fronted constituents to precede them within the periphery of the adverbial clause containing them. However, *as* and *though* seem to be the only two subordinating conjunctions which potentially allow this kind of fronting in English, e.g. in structures such as the following:

(4) a. (*Hard though* he tried —), he could not open the door
b. (*Try as* he might —), he could not open the door

By contrast, other subordinating conjunctions don’t allow periphery-internal fronting
— as we see from the ungrammaticality of the adverbial clauses bracketed below:

(5)  
  a. *[His passport because he had forgotten —], he had to return home  
  b. *[Harder if you try —], you will surely succeed  
  c. *[Hot when you feel —], you should take off your jacket

Thus it seems clear that while (some variant of) the analysis in (3) might be appropriate for though/as, it would not be appropriate for other subordinating conjunctions (like because, for example), since they do not allow this kind of fronting.

Still, one way of modifying the analysis in (3) in such a way as to account for because-type conjunctions not allowing fronting might be to suppose that they have a reduced periphery which contains only FINP. On this alternative view, in place of (3) above we would have (6) below:

(6)  They got wet [FINP [FIN because] it was raining]

An analysis along the lines of (6) would account both for the absence of fronting in because-clauses (since there is no position in the periphery above because to house the fronted constituent), and for absence of that in mainstream varieties (since there is no position in the periphery below FIN to accommodate that).

However, the claim embodied in analyses like (3), (6) that the subordinating conjunction is positioned in FIN in adverbial clauses which don’t contain that is undermined by the observation that subordinating conjunctions used without that can be followed by a range of peripheral constituents which are typically found in clauses containing FORCEP, including (as in the examples below) peripheral exclamative, interrogative, imperative, negative and topical constituents:

(7)  
  a. It’s easy to forget about Everton, [because what a good run they’re having!] (John Cross, Talksport Radio)  
  b. We puzzle over it a bit, and then brush it off and go on with our daily lives [because what can we do?] (steamcommunity.com)  
  c. I’m telling you [because please don’t let the rain deter you] (tripadvisor.com)  
  d. That is why I want a united Europe [because never again should we have wars amongst ourselves …] (otib.co.uk)  
  e. I’ve gotta see what I can do moving forward, [because the past, I can’t change] (Paul Stewart, Talksport Radio)  
  f. That’s [because Smalling and Jones, neither of them have made it for
Given the assumption that only FORCEP constituents can contain topicalised, focused or fronted constituents, it follows that adverbial clauses without that must contain a FORCEP projection. This conclusion is reinforced by the observation that the bracketed adverbial clause can have its own force, and hence be exclamative in (7a), interrogative in (7b), imperative in (7c), and declarative in (7d)-(7f).

What this suggests is that a co-ordinating conjunction like because is a causal subordinator which can have as its complement a clause which is interrogative, imperative, exclamative or declarative in force. A straightforward way of capturing this insight is to treat because as a SUB head which selects a FORCEP complement that can be interrogative, imperative, exclamative or declarative in type. On this view, the subordinate clause in (1a) above would have the peripheral structure shown below:

(8) They got wet [SUBP [SUB because] [FORCEP [FORCE ø/that] [FINP [FIN ø] it was raining]]]

In most varieties of English, the declarative FORCE head in a structure like (8) would receive a null spellout; but in a minority of varieties, it could be spelled out as that. However, this raises the question of why use of that in adverbial clauses should be allowed in some varieties but not in others.

One possible answer is that the use (or non-use) of that after a subordinating conjunction reflects a low-level difference in the PF spellout conditions for that. One implementation of this idea would be to suppose that minority varieties which allow SUB+that structures permit non-initial peripheral heads to be spelled out as that in appropriate kinds of embedded clause (including in clauses introduced by a subordinating conjunction/SUB) whereas mainstream varieties which don't allow SUB+that structures only permit a peripheral head to be spelled out as that when it is the highest head/first word in its containing clause. This would mean that most speakers will not allow the FORCE head in (8) to be spelled out as that because it is not the first word in the periphery of the adverbial clause containing it (the first word being the SUB constituent because), but a minority of speakers will allow such structures.

I set aside here what Liliane (in her 2012 monograph, ch.4) calls ‘central adverbial clauses’, which she argues to be closely integrated into the matrix clause, and to have a reduced peripheral structure.
Speakers of the minority variety which does not require that to be clause-initial would also be expected to allow the complementiser that to be used non-initially in other types of clause, like those below:

(9)  
   a. I think, [as expected, so far that the Greeks are not showing any real ambition] (Danny Mills, BBC Radio 5)  
   b. I think [Bayern Munich that they are a team to really watch in the final stages] (Andy Brassell, BBC Radio 5)  
   c. I just think [Arsenal, honestly, that they are gonna fall behind if they don’t sign a striker] (John Cross, Talk Sport Radio)  
   d. The fact that he’s been around for so long, I think, is testimony to [just what a good player that he is] (Tim Vickery, BBC Radio 5)  
   e. It’s unbelievable when you think what a big city Paris is and [what a potentially big club that PSG are] (Andy Brassell, BBC Radio 5)  
   f. It’ll be interesting to see [what sort of game that he puts on] (John Cross, Talk Sport Radio)  
   g. We’re hoping JLo will come into the studio, just so she can show us [what moves that they make] (Andy Goldstein, Talk Sport Radio)  
   h. I just don’t understand [why that you would present a fatted calf to sell] (John Cross, Talksport Radio)  
   i. I’m aware of the speed [with which that they work] (Tim Vickery, BBC Radio 5)

Here we find non-initial that preceded within the bracketted clauses by italicised peripheral constituents which include in situ adjuncts or dislocated topics, and fronted exclamative, interrogative or relative wh-constituents. As will be apparent, there would seem to be an overlap between speakers who allow use of that in adverbial clauses like those in (1), and speakers who allow that in the types of embedded clauses bracketed in (9): for example, Andy Brassell, John Cross, Andy Goldstein, and Tim Vickery use that in both types of structure. Such speakers do not restrict that to spelling out a periphery-initial head.

An alternative approach to accounting for the use of that in adverbial clauses would be to treat use or non-use of that as a matter of selection, and suppose that in standard varieties, subordinating conjunctions select a FORCEP complement with a null head, whereas in some minority varieties they can select a FOR-CEP complement whose head can either be spelled out as that, or receive a null spellout. However, this would raise questions about the nature of the relevant
selection restrictions. One possible answer would be to posit that an overt complementiser like that has to be licensed by an immediately adjacent superordinate constituent of an appropriate kind, and that standard varieties of present-day English impose the following constraint on complementiser licensing:

(10) Complementiser Licensing Constraint/CLC
In standard varieties of English, the complementiser that cannot be licenced by a peripheral head.

CLC would allow superordinate lexical heads like the verb think, the adjective sure and the noun claim to license use of that in complement clauses like those bracketed below:

(11) a. I think [that you are right]  
     b. I am sure [that he will come]  
     c. Your claim [that he lied] is preposterous

At the same time, CLC would rule out a structure such as that below, so preventing the interrogative complementisers whether/if from licensing that:

(12) I wonder [FORCEP [FORCE ø] [INTP Op [INT whether/if] [FINP [FIN ø/*that ] it will rain ]]]

More relevant to our present discussion is the observation that CLC would also rule out the use of that in adverbial clauses like that below (cf. 8 above) in standard varieties of English:

(13) They got wet [SUBP [SUB because] [FORCEP [FORCE that] [FINP [FIN ø] it was raining ]]]

This is because CLC bars that from being licensed by a peripheral head like the SUB(ordinating conjunction) because. By contrast, in minority varieties in which

\[We might try to derive the relevant selectional properties from independent properties of the conjunction and/or the complementiser. For example, in her 1992 CUP book Theory and Description in Generative Grammar: A Case Study of West Flemish, Liliane relates the obligatory of the complementiser dat/that in adverbial clauses (and other embedded clauses) in West Flemish to the observation that the complementiser inflects for agreement with the clause subject, and suggests that the complementiser has to be overt in order to spell out the agreement features which it carries. However, as Liliane herself notes (pc), such an analysis would be difficult to extend to English, given that the complementiser that in English is generally optional and does not inflect for agreement.\]
CLC is inoperative, structures like (13) are permitted. Indeed it would seem from examples like those below that structures like (12) are also permitted by some speakers:

(14) a. I just don’t know [whether that they will have the same attitude] (Mark Saggers, Talksport Radio)
    b. It’s just a question of [whether that Liverpool can get their money back] (John Cross, Talksport Radio)
    c. I’m not sure [whether that Spurs fans will accept him] (John Cross, Talksport Radio)
    d. I do wonder [whether that their squad lacks the depth of City’s] (Dominic Fyfield, Talksport Radio)
    e. England have enforced the follow-on. [Whether that they could have done it had it not been raining], I’m not sure (Jack Bannister, Talksport Radio)
    f. It’s not clear, though, [if that they’re just infecting the microbes that make us sick] (Carl Zimmer, BBC Radio 5)

However, speakers who treat that as inherently non-interrogative would be expected to reject sentences like (14), even if they accept the use of that in adverbial clauses like that bracketed in (13).

Clearly, there is much more to be said about the use of the complementiser that in adverbial clauses, but I shall not attempt to delve further into the matter here, since the point of this brief excursus into adverbial clauses is that it illustrates many of Liliane’s finest qualities. Firstly, she is an enormously productive and perceptive linguist: her book on Adverbial Clauses amply illustrates the depth of her scholarship, and the insights of her analysis. Secondly, she is always willing to help, whether by providing copies of articles which nobody else seemed to be able to find (including on one occasion an article by Jim McCloskey which Jim himself had lost!), or by delving into her archives of non-canonical structures in written English to find analogous structures, or by asking questions which nobody had previously thought to ask (e.g. when I was working on how come questions, she asked whether they allow Subject Drop, and that set me off on a new stream of thought). Enjoy your retirement, Liliane, free from the shackles of bungling bureaucrats, and from their mindless metrics which reduce evaluating the work of distinguished staff and their students to mind-numbing number-crunching.
Left peripheral NEG as a discourse particle

Henk van Riemsdijk

The present little article is in reality only a footnote. But it is a footnote to several topics. First it is a side remark on the use of sentence-initial negative markers in older versions of the Germanic languages, in particular Gothic, Old English and Old High German. Second, it shows that there are instances of sentence initial negative markers also in modern varieties of German, and furthermore that these cases are best treated as instances of discourse particles. These are all topics that I have never before worked on, but they do figure prominently in the work of Liliane Haegeman.

It is generally believed that sentence initial particles, which used to be quite frequent in older varieties of German, have been lost. See, for example, Petrova (2017), particularly section 4 entitled “The loss of left-peripheral sentence particles and the rise of modal particles in German”. In the few pages below I will argue that while there may not be many, there definitely are modal particles (or discourse particles) at the left edge in at least some varieties of present day German.

In some modern varieties of German, sentence initial occurrences of the negative particle (‘nicht’) can be observed. The phenomenon is primarily limited to Austrian German but is marginally also present (or at least perceived as ‘something that some people say’) in other southern varieties of German, see for instance the example (5) from Zurich German (Züritüütsch) below. The meaning of this sentence initial negative marker (henceforth NEG1) is, however, quite distinct from ‘normal’ sentence negation. NEG1 roughly expresses something rang-
ing from great surprise to indignation about the (non-negated) state of affairs described in the sentence. Here are some examples with various paraphrases:

(1) Nicht kommst Du jetzt mit noch einer Fussnote!
   not come you now with yet another footnote
   'Don’t tell me you are coming up with yet another footnote!'
   'It just cannot be true that you are presenting another footnote!'

(2) Nicht habt ihr die Kekse schon alle gegessen!
   not have you the cookies already all eaten
   'Don’t tell me that you have eaten all the cookies!’
   'I cannot believe that you have eaten all the cookies.’
   'You have already eaten all the cookies? – You must be kidding!’

Sometimes there is a similar alternative with NEG1 followed by a that-sentence:

(3) a. Nicht kommt mir der Kerl noch einmal ins Haus!
   not comes me that guy another time in the house
   'It just is not going to happen that that guy comes into my house
   again!’
   b. Nicht dass mir der Kerl noch einmal ins Haus kommt!

Note that (1), (2) and (3a) have V2 and not VFinal as in (3b). This is reminiscent of Swiss German wäisch (‘you know/know you’), cf. van Riemsdijk (2001b), which I called a wh-prefix. Here’s an example:

(4) a. Wäisch wann de Hans geschter häi choo isch?
   do-you-know when the Hans yesterday home come has
   ‘Do you know when Hans came home yesterday?’
   b. Wäisch wann isch de Hans geschter häi choo?
   know-you when has the Hans yesterday home come
   ‘You won’t believe what time Hans came home yesterday!’

(4a) is a real question, but (4b) is something like a rhetorical question: the speaker knows the answer and is so horrified by the answer (sc. 4am) that he wants to tell the listener. (4a) has a wh-question embedded under the question introducing verb wäisch (‘do you know’), which is also a yes-no question. Accordingly the embedded verb is in the final position. In (4b) the wh-question is the main clause (V2) in which wäisch is prefix to the wh-word, comparable to wh-suffixes like the
The NEG1 examples, while predominantly found in Austria do not seem completely out in other German-speaking areas. Here is a Swiss German example I find acceptable:

(5) Nöd häsch scho alles ggässe.
    not have-you already everything eaten
    ‘Don’t tell me you have already eaten everything!’

But for me the VFinal version is ungrammatical.

(6) *Nöd dass du scho alles ggässe häsch.

There are other variants that are clearly similar but have a different use, definitely not as discourse particles, and are, as far as I can tell, much less common. The Austrian daily Der Standard, for example, has a column called

(7) Ned sei deppert!
    not be stupid
    ‘Don’t be stupid!’

The question arises as to what the origin of NEG1 construction is. Old English did have initial negation, as the following example shows:

(8) Nat may the woeful spirit in my heart declare one bit of all my sorwes smerte sorrows painful
    ‘The woeful spirit in my heart cannot clarify one bit of all my painful sorrows’
    (ca. 1385, Chaucer CT.Kn. A.2765: MED)

But examples of this sort can also be found in Old High German, as shown in the following examples from Migdalski (2016: 117, ex. (53c); 120 ex.(60)):

(9) Ni liugu ih dauid.
    NEG lie I David
    ‘I will not lie to David’

\[\text{2 See van Riemsdijk (2001b). I am, of course, tempted to consider the prefix wäisch a graft, just like, e.g. far from in a far from simple question, cf. van Riemsdijk (2001a).}\]

\[\text{3 The example is taken from Koike (2016: 251, ex. (170a)).}\]
(10) Ni-santa got sinan sun.
   ‘God did not send his son’ (OHG, Tatian, 407, 30, (Axel-Tober 2007: 61))

Breitbarth (2013) lists also some examples of initial negation from Old Low German (her example (1a): Old Saxon) and Old Dutch (her example (2a): Old Low Franconian).

(11) ‘Ni ĖCi, quað he, ‘that barn godes…’
   NEG am I spoke he the child god.GEN
   ‘I am not the child of God, he said’ (Old Saxon. Heliand, 915)

(12) Ne farlātu tu mi!
   NEG forsake you me
   ‘Do not forsake me!’ (Old Low Franconian. WP LXX.9.2)

Other, later, examples are found in certain Bible or Torah translations of commandments. Here are a few cases I found:

(13) Aber wenn ist an ihm ein Makel, lahm oder blind, irgendeinen
    but if is on it a flaw lame or blind, any
    schlechten Makel, nicht sollst Du es schlachten für JHWH, deinen
    bad flaw not should you it slaughter for JHWH, your
    Gott.
    God
    ‘But if it has a flaw, being lame or blind, any bad flaw, do not slaughter it for JHWH, your God!’

(14) Nicht sollst du ehebrechen, nicht sollst du töten, nicht sollst
    not should you commit-adultery, not should you kill, not should
    du begehren, etc.
    you covet
    ‘Not shalt thou commit adultery, not shalt thou kill, not shalt thou covet,
    …’ (Romans chapter 13)

(15) Nicht darf er sich laben an Bächen, flutenden Strömen von Honig
    not may he himself refresh at creeks, flowing rivers of honey
    und Milch.
    and milk

The examples in (13) to (16) are from Paganini (2009), example (17) is from Krüger (2001). The glosses and translations are mine.
Aber die Nichtswürdigen, wie geflohene (d.h. verabscheute) Dornen sind sie alle, denn nicht werden sie mit der Hand gefasst.

(2. Samuel chapter 23 v. 6-13; Luther translation)

Viele wirst Du lenken, wenn die Vernunft Dich lenkt: von ihr wirst du lernen, was und wie du es anpacken musst: nicht wirst du den Dingen anheimfallen.

(Krüger 2001: Seneca Epistles 37,4)

Turning now to the syntactic position and semantic role of NEG1, observe first, that it’s interpretation is, essentially, extraclausal. This is easily seen from the ample translations provided above for the examples in (1)-(3). This observation is further strengthened by the fact that the NEG1 element is itself often paraphrased by some short sentence such as ‘don’t tell me’ or ‘it just cannot be true’ in example (1). The special status of NEG1 is also evidenced by the fact that the negative marker induces a clear meaning contrast with the clause internal negation. Consider the following triple of examples:

(18) a. Nicht kommst Du morgen wieder zu spät.
   ‘Don’t tell me you are going to be late again tomorrow!’

b. Du wirst morgen nicht wieder zu spät kommen.
   ‘You will not be late again tomorrow!’

c. Du wirst doch nicht etwa morgen wieder zu spät kommen?
   ‘Surely you will not be late again tomorrow – don’t you dare!’

(18a) is the typical case of NEG1: the speaker expresses great dismay if the hearer were to repeat being late tomorrow. It is a thinly veiled threat of blame or even punishment. (18b) is a bland statement: you will not be late tomorrow. Of course
it could, given the right pragmatic circumstances be used as dismay or threat, but it can just as easily mean something like it will not be snowing tomorrow, so the train will be on time and there is no reason why you should be late. However, shows that standard discourse particles (or modal particles as they are sometimes also called) such as *doch* and *etwa* can achieve an effect quite similar to the one caused by *ēĊČ*.

In Bayer & Struckmeier (2017), one of the properties of discourse particles they stress is that they are similar to 'high adverbs' (see in particular Cinque 1999). There are indeed good reasons for equating *ēĊČ* with high adverbs, however I will not venture into the morass of the complex structure of the left periphery. See Rizzi (1997) for influential proposals and Haegeman & Lohndal (2017) for some critical discussion.

In view of the fact that you cannot be higher than the leftmost element in a sentence, it is, therefore, important to ask if there are any other potential cases of leftmost adverbs, under which I would subsume adjectives that are neither attributive nor predicative in relation to some element within the sentence. Martin Prinzhorn (p.c.) has pointed out to me a term paper by a student of his, David Diem, who has observed that a relevant construction can be found in Swiss German with certain adjectives, in particular *schön* ('nice'), *guet* ('good') and *schad* ('shame, pity, too bad'). Here are some examples that Diem found in the Swiss SMS Corpus (Stark et al. 2009-2015).

(19) Schön gits di min schatz! 'Nice that you exist, my dear!'
(20) Guet bisch früe ufs fäscht choo! 'A good thing you came early to the party!'
(21) Schad häsch du morn nöd frei,... 'A pity you are not free (sc. from work) tomorrow…'

One possible hypothesis might be that the adjective is really a predicate adjective in a structure like 'it is A', in Swiss German 's isch A', with the 'it is' part elided (or silent). Indeed such predicative adjective constructions can be constructed with either V2 or with complementizer and VFinal. The VFinal variant of (21), for example, would be (22).
Recall that we observed above that the NEG1 is often best paraphrased by a sentential expression as was mentioned between the examples (17) and (18) above. Future research will have to show whether left-peripheral discourse particles are limited to these instances of reduced mini-clauses.

To conclude, here is a particularly nice example of a Swiss German adjectival adverb in first position.

(23) Schön isch-s hüt schön!
     nice is-it today nice
     'It’s nice that it’s nice today!' (= It is a pleasant fact that the weather is nice today)

References


‘Recycling’ evidentiality: a research program

Johan Rooryck

Abstract

In this paper, I challenge the idea that evidentiality constitutes a grammatical category of its own. I propose that it should be viewed as a grammatical mechanism that creates evidential meanings by recombinations of features ‘recycled’ from other, more basic grammatical categories.

1 Evidentiality as a grammatical category

The languages of the world display a wide variety of grammatical markers to express evidentiality, ranging from dedicated morphemes to adverbs and parentheticals. Despite this expressive variety, evidentials form a closed class system of mutually exclusive markers. The range of evidential meanings that is expressed is limited to 3 or 4 in most languages. There is also a striking consistency across languages in the types of evidence expressed by these markers. Many authors have proposed a basic distinction between direct and indirect evidence types (see Table 1) which can be further subdivided into at least hearsay and quotative on the one hand, and inference/conjecture on the other (Givón 1982, Bybee 1985, Willett 1988: 57):

<table>
<thead>
<tr>
<th>Direct evidence</th>
<th>Personal experience of the speaker: visual/auditory/other sensory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect evidence</td>
<td>Reported to the speaker: hearsay/quotative</td>
</tr>
<tr>
<td></td>
<td>Inference/conjecture by the speaker.</td>
</tr>
</tbody>
</table>

I dedicate this article to Liliane Haegeman on the occasion of her retirement, with respect, admiration, and gratitude for her impressive work in linguistics.
These evidence types can be illustrated by Faller’s (2002) data on Cuzco Quechua in (1), a language that features dedicated morphemes for three types of evidential meaning:

(1)  
\begin{enumerate}
\item Para-sha-n-\textit{mi}  
\begin{itemize}
\item rain-\textit{PROG-3-}\textit{BPG}
\item p=’It is raining.’
\item EV: speaker sees that it is raining
\end{itemize}
\item Para-sha-n-\textit{chá}.
\begin{itemize}
\item rain-\textit{PROG-3-}\textit{CONJ}
\item p=’It is possibly raining.’
\item EV: speaker conjectures that it might be raining.
\end{itemize}
\item Para-sha-n-\textit{si}.
\begin{itemize}
\item rain-\textit{PROG-3-}\textit{REP}
\item p=’It is raining.’
\item EV: speaker was told that it is raining
\end{itemize}
\end{enumerate}

According to Faller (2002), the morpheme -\textit{mi} in (1a) indicates that the speaker claims to have direct, visual evidence (BPG = best possible grounds) for the fact that it is raining. By contrast, the morpheme -\textit{chá} in (1b) expresses the speaker’s conjecture, and -\textit{si} in (1c) reflects hearsay (reportative) (Faller 2002).

The organization of evidence types in Table 1 suggests that evidentials constitute a cross-linguistically homogeneous grammatical category. The evidence types are not only constant across languages, they are also limited in terms of their semantics. In principle, the number of evidence types could have been much higher. As Speas (2008) notes, it is easy to imagine additional indirect evidence types, such as divine revelation, custom, legal edict, or parental authority. No language has such indirect evidentials: there are just the subcategories of direct and indirect evidence, with indirect evidence dividing further into hearsay/quotation and inference/conjecture. The limited nature of this set of subcategories is then implicitly taken as evidence for the idea that evidentials form a closed class of mutually exclusive functional elements that exhibit a paradigmatic pattern of oppositions, just like other grammatical categories. The data from Cuzco Quechua, where dedicated evidential morphemes function as mutually exclusive elements of a closed class paradigm, seem to confirm such an analysis.
2 ‘Splitters’ and ‘recyclers’

A major discussion regarding evidentiality revolves around the question whether evidentiality should be ‘split off’ as a grammatical category in its own right, or whether it is a grammatical notion that ‘recycles’ existing grammatical categories for its own purpose. ‘Splitters’ argue that the specific range of evidential meanings in paradigmatic opposition militates in favor of a separate grammatical status for evidentiality, on a par with categories such as tense, aspect, modality, or person. Evidence types such as witness, hearsay (reportative), and inference are then viewed as universal, elementary features of human language. Just like other functional categories, evidentiality is assigned a specific functional head in the left periphery of the functional domain of the clause (Cinque 1999, Rooryck 2001a,b). The view that evidentiality is a full-fledged grammatical category in its own right is implicitly or explicitly adopted by linguists of various theoretical persuasions (Hardman 1986, Cinque 1999, de Haan 1999, Lazard 2001, Aikhenvald 2004, Davis et al. 2007, Aikhenvald 2018). It is fair to say that it reflects the dominant perspective.

‘Recyclers’, by contrast, seek to relate evidential meanings to the interaction of more elementary grammatical categories, such as tense, aspect, modality, sentence-type, and person. In many languages, evidential meanings do not involve dedicated morphemes at all. Rather, specific evidence types are parasitic on tense, aspect, modality, sentence-type, and person. In these languages, evidential meanings can be analyzed as the result of ‘recycling’ these more basic categories. In what follows, I will reinterpret analyses of evidentiality that relate it to other categories in terms of ‘recycling’. Of course, this does not necessarily mean that the authors of these analyses see themselves as ‘recyclers’ in the sense that I am proposing here.

---

2 This article owes an important intellectual debt to Peggy Speas’s work on evidentiality in the last 15 years. To a large extent, the present article does not do much more than taking further some of Peggy’s positions and radicalizing them.

3 See Speas (2008) for a critical analysis of this position.

4 The idea of ‘recycling’ as a general grammatical mechanism has recently been given a more solid theoretical footing by Biberauer (2017) as an instance of her Maximize Minimal Means (MMM) principle.
3 Two sources for reportatives

Let me illustrate this with the example of reportative evidentiality. It is well known that the German modal *sollen* ‘should’ does double duty as a reportative evidential (Schenker 2008, Faller 2017):

(2) Hubert soll in Berlin sein

Hubert should in Berlin be

i. ‘Hubert should be in Berlin (given his obligations)’ (deontic modal)

ii. ‘Hubert is said to be in Berlin’ (reportative) [German]

Faller (2017) analyses reportative *sollen* ‘should’ in terms of extending the modal base of deontic *sollen* ‘should’ to an information modal base. This change in the modal base can be interpreted as an instance of the reportative ‘recycling’ the deontic modal for its own purposes. Other languages recycle different categories to build reportative meaning. Bruil (2014, 2015) shows that in Ecuadorean Siona, reportative evidentiality is part of the clause-typing system of the language, alongside declarative and interrogative:

(3) a. Aibi něcaji.
   Ai-i-bi nĩ’hika-hi. (declarative)
   old-NCL:M-SBJ stand-3S.M.PRS.ASS
   ‘The old man is standing.’ (Bruil 2014: ch6, 11a)

b. Aibi něcaquē?
   Ai-i-bi nĩ’hika-kī? (interrogative)
   old-NCL:M-SBJ stand-2/3S.M.PRS.N.ASS
   ‘Is the old man standing?’ (Bruil 2014 ch6, 11b)

c. Jao ti co’meco beocoña.
   Hā-ō ti ko’mē-ko beo-ko-jā. (reportative)
   dem.dst-ncl:f ana row-NOM.S.F neg.exist-2/3S.F.PRS.N.ASS-REP
   ‘She doesn’t have gas.’ (I am told). (Bruil 2014 ch6, 19) [Ecuadorean Siona]

Bruil (2014) analyses this integration of the reportative in the clause-typing system in the following way. She observes that in the declarative and the interrogative, epistemic authority for the proposition lies respectively with the Speaker and the Addressee. The Ecuadorean Siona reportative fits into this system by combining epistemic authority with a non-speech participant. Epistemic authority for the proposition is attributed to a third party that is neither Speaker nor Ad-
dressee. In this way, the essential meaning of reportative is derived (Bruil 2014). Importantly, the reportative functions entirely within the paradigmatic oppositions afforded by the grammatical category of sentence-type, rather than as a subcategory of evidentiality in the sense of Table 1.

The interest of the examples from German and Ecuadorean Siona lies in the fact that reportative meaning in either language is built from very different ingredients: in German, the reportative recycles deontic modality, while in Ecuadorean Siona, it is parasitic on the clause-typing system. These examples suggest that it is unlikely that ‘reportative’ is a primitive of the grammar, i.e. a particular paradigmatic exponent of an overarching and independent grammatical category ‘evidentiality’. Rather, it suggests that such evidential meanings arise from recombinations of more abstract ingredients that are made available by more basic grammatical categories.

4 More examples of evidential ‘recycling’: tense and aspect

The reportative is not the only evidential meaning that can be derived from more basic grammatical categories. Nikolaeva (1999) observes that evidentials in Ostyak (Finno-Ugric, Siberia) interact with present and past tense. She derives the evidential meanings of hearsay and inference in terms of equivalence or overlap between situation types. Chung (2005) shows that Korean evidentials are homophonous with aspect and mood morphemes, and develops an analysis in which these interact with the location of the (1st person) speaker to derive evidential meanings. Lau & Rooryck (2017) argue that abstract properties of event structure, more specifically the event stages in accomplishments and achievements, are recycled in terms of information stages for the purposes of evidentiality. A semantic analysis in terms of stages allows them to bring out the close relation between aspect, indirect evidentiality, and mirativity in languages from Turkish and Bulgarian to Washo (Hokan, USA) and Hare (Athapaskan, Canada). For instance, the Turkish morpheme -miş can express perfect aspect, indirect evidentiality, and mirativity:

(4) Kemal gel-miş [Turkish]
    Kemal come-PERF
    ‘Kemal came.’

(a) INFERENCE: The Speaker sees Kemal’s coat hanging in the front hall,
but has not yet seen Kemal.
(b) HEARSAY: The Speaker has been told that Kemal has arrived, but has not yet seen Kemal.
(c) SURPRISE: The Speaker hears someone approach, opens the door, and sees Kemal—a totally unexpected visitor.”

In terms of Lau & Rooryck (2017)’s analysis, this is possible because -miş is underspecified in terms of the nature of the stages involved. Informally put, -miş does not care whether its stages are of the event type (aspect) or of the information type (evidentiality/ mirativity). The morpheme -miş is primarily sensitive to the presence of a final stage holding at utterance time, regardless of its status as an information stage or an event stage. Once again, it looks like the more basic category of aspect is ‘recycled’ – bleached in this case – to express evidential meanings.

Such analyses suggest that evidentiality is not a grammatical category on a par with tense, aspect, modality, or sentence-type, but that it rather is a grammatical mechanism that is able to exploit these more primitive categories for its own purposes. This observation makes it very unlikely that evidentiality should be assigned its own functional category in the left periphery, in the sense of Cinque (1999). For all intents and purposes, it is probably sufficient to syntactically represent Speaker and Addressee in the left periphery (Garrett 2001, Speas & Tenny 2003, Haegeman & Hill 2013), since speech participants play an important role in ‘building’ evidential meanings. Note that Speaker and Addressee in this sense are just special instances of the more abstract features 1P and 2P.

Following Rooryck (2001a,b) and Speas (2004b,a) among many others, I would argue that syntactically represented speech participants are part and parcel of the elementary syntactic categories that are ‘recycled’ by evidentiality.

The idea that evidentiality does not occupy a well-defined, single position in the functional domain as other grammatical categories do is corroborated by Blain & Déchaine (2007). On the basis of various dialects of Cree, Blain & Déchaine (2007) argue that evidentials differ from each other as a function of the syntactic domain where they are introduced (their Evidential Domain Hypothesis). While this analysis does not call into question evidentiality as a separate category, it suggests that evidentiality arises at various levels of the functional spine, and need not be assigned its own morphosyntactic representation.

5 Speas (2008) also questions the evidence for an evidential head. However, she does argue in favor of evidentiality as a distinct functional category on a par with tense (see Speas 2010).
5 Evidentiality as a closed class?

The analysis of evidentiality as a distinct grammatical category is further called into question by the fact that its boundaries are not very well established. More precisely, it does not constitute as much of a closed class with well-defined paradigmatic oppositions as the example of Quechua in (1) and the categorization in Table 1 would suggest. There are many instances of vaguely related and overlapping notions that are hard to reduce to strong paradigmatic oppositions. Even in Table 1, the distinction between inference and conjecture is not an immediately obvious one. Anderson (1986) provides a map of mental space for evidential meanings that includes different types of inferentials, as well as a category of expectation (as in English to be supposed to). In Carib, Hoff (1986) argues that introspective evidentials (knowledge through inference) should be distinguished from extraspective evidentials (culturally available knowledge).

More generally, the distinction between evidentiality and modality is notoriously difficult to establish (see Speas (2008, 2018) for an overview). Note also the overlap between indirect evidentiality and mirativity in Turkish and other languages mentioned above in the discussion of (4). The notion of mirativity does not even occur in Table 1, while it can be linked to indirect evidentiality. Building on earlier work by Adelaar (1977, 2013) on mirativity in Quechua, Mexas (2016) argues that mirativity should be understood as ‘sudden realization’: a punctual transition from the epistemic state of ignorance to that of awareness. That redefinition brings mirativity much closer to the indirect evidentiality expressed by inference. As Lau & Rooryck (2017) argue, inference can be viewed as a non-punctual process towards an epistemic transition, and thus only differs in terms of the stages leading to the transition. However, this difference between mirativity and inference is in fact not a properly evidential distinction, but one that is based on notions of the succession of stages that are originally aspectual in nature. Although further study is needed, it is my contention that all oppositions in the evidential domain can ultimately be reduced to properties that are provided by other, more basic grammatical categories.

Assuming that evidential meanings are indeed based on ‘recycled’ features from more basic grammatical categories, the question arises how languages with designated evidential morphemes, such as Quechua, should be dealt with. This is an important question, since languages in which dedicated evidential

---

6 Mirativity is sometimes argued to be entirely distinct from evidentiality (see Hengeveld & Olbertz, 2012; Aikhenvald, 2012).
morphemes are obligatorily expressed are often set apart from languages without such evidentials (Aikhenvald 2004, 2018). My proposal here would be that dedicated evidential morphemes are the result of grammaticalizing the properties of more basic grammatical categories into specific markers. de Haan (1999) shows that the diachronic sources of evidentials involve verbs of speech, vision, and inference; spatial expressions, and tense and aspect. However, the ‘standalone’ nature of a set of dedicated evidential morphemes is not enough to provide them with their own dedicated grammatical label. Rather, they are often morphosyntactically dependent on other categories in the functional domain. It is well known that the exact morphosyntactic position of dedicated evidential morphemes differs widely across languages, a point also made by Blain & Déchaîne (2007). This positional variability is another argument against assigning them a specific position and label in the functional domain.

Despite the fuzziness of a number of evidential meanings, the fact remains that evidential morphemes often appear to function as if they were part of a closed class as in Table 1. However, from the perspective of ‘recycling’, this apparent paradigmatic organisation should be viewed as a consequence of the limited number of categories that are available for evidential ‘recycling’. A limited variety of recycled categories can yield the circumscribed set of evidential meanings illustrated in Table 1. What remains surprising is that ingredients that are so different at the outset result in the same evidential outcomes, as in the case of reportatives discussed in section 3 above. A research program into the ‘recycling’ nature of evidentiality should explore the constraints on the type of categories that can be recycled for evidential purposes. The categories that can be recycled must have a set of abstract features that are compatible with evidential meanings. In the next section, I will examine Speas’ (2010) proposal for such an abstract analysis of evidentiality, evaluate to what extent it fits ‘recycling’ purposes, and formulate a proposal of my own that complements Speas’ (2010) approach.

6 Analyzing the basic features behind evidentiality: Speas (2010)

Speas (2010) develops a theoretical approach to evidentiality that aims at deriving specific evidential meanings from more abstract features and mechanisms. She argues against the notion of ‘evidence’ as a grammatical primitive, and pro-
poses an analysis of direct and indirect evidentiality in terms of the accessibility of situations, much inspired by Kratzer's (1977) analysis of modals in terms of a modal base of accessible situations, and a Reichenbachian approach to tense. Speas (2010) proposes three types of situation:

(5)  
  a. Evaluated Situation (ES): The situation of which p is true  
  b. Reference Situation (RS): A situation or set of assumptions relevant to ES and DS  
  c. Discourse Situation (DS): The situation in which the sentence is being uttered

Between these Situations, two types of relations can hold: inclusion or accessibility. Indirect evidentiality involves a context in which the Reference Situation is accessible from the Evaluated Situation, while in direct evidentiality, the Reference Situation includes the Evaluated Situation. Further distinctions within direct and indirect evidentiality are made possible by the notion of Discourse Situation. In Hearsay, RS is not only accessible from ES, but RS is in turn accessible from DS. In other words, an ES in which ‘Kim saw a bear’ is true, is accessible to an RS where Kim tells the speaker that she saw a bear, and this RS is in turn accessible to the Discourse Situation where the speaker tells the hearer ‘Kim saw a bear’, while including into that statement an evidential expressing that this information was made accessible through Kim’s telling the speaker. By contrast, in indirect evidentials (inferences), RS includes DS.

Direct evidentials can likewise be further divided into ‘internal’ situations only the speaker can know (personal experience), and situations external to the speaker that are directly perceived through other senses. This is for instance the case in Eastern Pomo:

(6)  
  a. bi.Yá pha.bé-kh-ink’e  
     hand burn-PUncTUAL-SENSORY  
     ‘I burned my hand’ (I feel the sensation of burning in my hand)  
  b. mí.-p-al pha.bé-k-a  
     3.sg.-male-PATIENT burn-PUncTUAL-DIRECT  
     ‘He got burned’ (I have direct evidence, e.g., I saw it happen)[Eastern Pomo]

Speas (2010) analyses personal experience as a context in which RS includes both ES and DS. By contrast, direct evidentiality through other sensory perception is defined as a context in which RS not only includes ES, but is also accessible from
DS. In other words, Speas (2010) reanalyzes the labels for evidential meanings in terms of the relations between the Evaluation Situation and a Reference Situation, and the Reference Situation and the Discourse Situation.

| “Personal experience” | RS includes ES  
|                       | RS includes DS  
| “Direct”              | RS includes ES  
|                       | RS is accessible from DS  
| “Indirect”            | RS is accessible from ES  
|                       | RS includes DS  
| “Hearsay”             | RS is accessible from ES  
|                       | RS is accessible from DS  

(7)

The aim of this analysis is of course very close to the one I am trying to defend here: evidentiality is analyzed in terms of more abstract features and relations that are shared with other categories. As Speas (2010) notes, the notions of Evaluation, Reference, and Discourse situation are very close to the notions Event time, Reference time and Discourse time from tense and aspect. Similarly, the inclusion and accessibility relations are very close to inclusion and precedence in tense and aspect. This closeness allows for a better understanding of the many cases in which tense and aspect morphemes are 'recycled' with evidential meanings.

However, Speas (2010) analysis also has a number of drawbacks. Just like in Reichenbachian approaches to tense and aspect, many more relations between RS, ES, and DS are possible in principle than those described in (7), and it is not clear why these are excluded or unattested. In (7), RS always enters an inclusion or accessibility relation with either DS or ES, but this asymmetry does not seem to derive from anything within the evidential system. It is also not clear whether indirect (inference) and hearsay are characterized with sufficient detail in this system. Speas (2010) states that the inference relation is characterized by an inclusion relation of DS in RS, because when making an inference, a speaker takes into account the current state of affairs. While that may be true, this inclusion does not capture the essence of what an inference is. An inference is a process of deduction on behalf of the speaker, who considers various types of contextual information to draw the conclusion that is likely that something is the case. The description of inference in terms of accessibility and inclusion does not take into
consideration that inferentials are always related to the speaker: the inference must be drawn by the speaker. Nothing in this description precludes an inferential evidential that would express an inference made by a third party, with the speaker simply reporting that this counts as the knowledge source of the statement. In other words, the description of inference in terms of accessibility and inclusion is compatible with the notion of inference, but it does not describe it precisely enough. Similarly, the description for hearsay does not include reference to the fact that the source of the hearsay always must be a third party. In other words, there are no hearsay evidentials dedicated to information provided by the hearer. Again, this does not quite follow from the description in \[7\].

7 Evidentiality and the proximal–distal distinction

However, I do believe that the analysis in \[7\] is on the right track, and rather than replace it I would like to propose an alternative perspective on the direct/ indirect evidentiality distinction that constrains Speas’s (2010) approach further. My proposal will perhaps also make it easier to evaluate which criteria basic grammatical categories have to fulfill to be able to be recycled into evidential meanings.

I will first examine the distinction between direct and indirect evidentiality in Table 1. Direct evidentiality signals that the speaker was present at the event, since they experienced the event by seeing, hearing, or even smelling it. Indirect evidentiality, whether by hearsay or inference, signals that the speaker was not present at the event. This direct/ indirect distinction closely resembles the universally attested spatial distinction between proximate and distal, a distinction that can be observed in many grammatical categories. As is well known, the proximate–distal distinction differentiates what is spatially close to the speaker to what is further away from the speaker. The distinction also often separates the visible from the invisible. This spatial distinction appears most clearly in the opposition between the English demonstratives this and here (proximate), and that and there (distal) (Diessel 2014). Proximal and distal demonstratives are used to establish a joint focus of attention between speaker and hearer (Diessel 1999).

This is of course also what evidentials do: they establish a joint focus of attention between speaker and hearer on the way in which the speaker obtained the information expressed in the statement presented to the hearer. Direct evidentiality can be viewed as proximal: it indicates that the event described in the
proposition was close to the speaker, and thus sensorily accessible (e.g. visible).\footnote{There are many antecedents for the relation between direct/ indirect evidentiality and deixis. Schlichter (1986) treats the Wintu evidential system in terms of a deictic opposition between unmarked visible events versus nonvisible events. Garrett (2001) argues that the direct evidential in Tibetan is a demonstrative assertion marker. Chung (2005) refers to 'spatial deictic tense' to derive direct evidentiality.}

Indirect evidentiality is distal: it indicates that the event was far away from the speaker, and sensorily inaccessible (hence often invisible). An indirect/distal evidential thus implies that the speaker indirectly obtained information about the event. So while this and that oppose proximal and distal objects, and here and there proximal and distal locations, direct and indirect evidentials can be characterized as referring to proximal and distal events. Under this view, the linguistic representation of knowledge and truth has a spatial basis. However, this spatial nature should not be viewed in terms of physical location but in terms of how events are presented as directly (proximal) or indirectly (distal) accessible to the speaker.\footnote{Note that even demonstratives do not always refer to deictic locations, as in expressions like There was this man I knew, where proximal this does not refer to a person who is locatively or contextually close to the Speaker.}

This entails that markers for direct/proximal evidentiality can be derived from verbs of seeing, since what is visible to the speaker is proximal in the relevant evidential sense.

The proximate–distal distinction is also fundamental in nonlinguistic cognitive capacities such as spatial navigation (e.g. Tommasi et al. 2012). A similar distinction is present in the two core cognitive systems for geometry distinguished by Spelke et al. (2010): a system for analyzing nearby visual forms is opposed to a system for navigating larger spaces. In other words, the way in which humans present the reliability and source of their statements may reflect a deep-seated capacity shared with cognitive systems beyond language. I believe this connection needs to be further explored to better understand the relation between intra- and extralinguistic constraints on linguistic representations.

This program is in line with the ‘three factors’ model for language of Chomsky (2005), who proposes that the faculty of language is composed of (a) a universal blueprint for language (Universal Grammar), (b) experience and usage, and (c) general cognitive factors. The proximal–distal distinction may well be such a general cognitive factor, and the exact nature of its interaction with language remains to be investigated.

Linguistically, this approach makes strong predictions for the way in which evidentials develop diachronically across languages, and for their interaction
with other categories: only grammatical categories that are compatible with a proximate–distal distinction can be recycled as direct–indirect evidentials. This strong ‘recycling’ hypothesis of course needs to be tested against a wide array of languages with evidential systems.

In addition to the proximal–distal dimension that translates the direct–indirect distinction in evidentiality, I would like to discuss another dimension that crosscuts the first one, and allows to describe the same four-way distinction that Speas (2010) had in mind. As mentioned before direct evidentiality can be further subdivided into situations that are only known to the speaker (something the speaker feels), and situations that are external to the speaker but directly perceived by them. Direct evidentials can be speaker-internal or speaker-external. Interestingly, Lau & Rooryck (2017) make a similar reference to the internal–external distinction in their analysis of inference and hearsay as directly related to the Aktionsart category of accomplishment. They analyze hearsay as an evidential with which the Speaker signals that the information presented in a proposition p was arrived at by a Speaker-external information update process of a source external to the speech situation informing them that p. By contrast, inferentials are minimally different from hearsay in that they refer to a Speaker-internal information update process using a variety of sources available to them. Evidentials expressing inference involve a Speaker-internal mental process of gradual ‘building up’ of the information that culminates in the Speaker possessing the relevant information expressed in p. With hearsay evidentials, the speaker is the recipient of external information, and is therefore positioned at the final stage of a process of transfer of information initiated by a third party. In both hearsay and inference, the speaker lacks direct access to the information.

The dimensions of proximal/ distal and Speaker-internal/ Speaker-external are sufficient to describe the same set of evidentials as those distinguished by Speas (2010), as shown in the following table:

<table>
<thead>
<tr>
<th>Speaker-access</th>
<th>Proximal situation</th>
<th>Distal situation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker-internal</td>
<td>Personal experience</td>
<td>Inference</td>
</tr>
<tr>
<td>Speaker-external</td>
<td>Sensory</td>
<td>Hearsay</td>
</tr>
</tbody>
</table>

Note that both of these dimensions are ultimately spatial in nature. While the proximal–distal distinction is well anchored in various other grammatical categories and processes, the nature of the Speaker-internal vs Speaker-external
distinction seems much less clearly related to existing categories, and requires more investigation. In any case, I believe that this ‘spatial’ analysis of the abstract features underlying evidential categories provide a concrete set of criteria that more basic categories have to comply with in order to be eligible for their ‘recycling’ as evidentials.

8 Extending the Speaker internal/external distinction

The Speaker-internal/-external dimension may also allow for a new understanding of the relation of evidentiality with egophoricity and allophoricity (see Tour- nadoré 1991, Hargreaves 2005, Widmer 2017 and references cited therein). Following Hargreaves (2005), Widmer (2017) views the egophoric/allophoric distinction as one that distinguishes between privileged vs non-privileged access to knowledge of a situation. On the basis of data from Tibeto-Burman languages, Widmer (2017) distinguishes egophoric markers that express ‘actional involvement’ of the speaker from from egophoric markers that express ‘epistemic involvement’. Egophoric markers expressing actional involvement mark the Speaker as a participant in a situation, while egophoric markers of epistemic involvement only indicate knowledge or experience of a situation.

This distinction is very similar to the one I made above between Speaker-internal and Speaker-external access to a situation. In other words, the actional/epistemic involvement distinction within egophoricity looks like another recombination of more basic primitives of the grammar. The notion of ‘privileged access’ does not seem to involve the proximal/distal distinction. By contrast, the allophoric (non-privileged access) system does seem to interact with evidential markers: allophoric markers can also express direct and inferential evidentiality in Bunan (Widmer 2017: table 2). In terms of the system developed in (8) above, this would mean that allophoricity does make use of the proximal/distal distinction as applied to situations. Widmer’s (2017) observations about Bunan could therefore be reinterpreted along the lines of the table in (9):

<table>
<thead>
<tr>
<th></th>
<th>Privileged access (egophoric)</th>
<th>Non-privileged access (allophoric)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>situational involvement</td>
<td>proximal situation</td>
</tr>
<tr>
<td>Speaker-internal access</td>
<td>actional involvement</td>
<td>proximal situation</td>
</tr>
<tr>
<td>Speaker-external access</td>
<td>epistemic involvement</td>
<td>distal situation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Direct</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inference</td>
</tr>
</tbody>
</table>
Once again, I hope to have shown that the relations between these various notions can be rethought in terms of different primitives: privileged access/egophoricity can be viewed as a non-spatial type of access of the Speaker to a situation.

9 Quotatives and reportatives

The table in (8) refers to hearsay as a context that involves speaker-external access to a distal situation. Throughout this paper, I have used the terms hearsay and reportative interchangeably, but I have not discussed quotatives. Note however that Table 1 puts hearsay and quotative on the same level, as part of indirect evidence. In terms of the table in (8), that cannot be quite right: quotatives are markers that attribute a sentence to a different speaker, and involve quoted speech. By contrast, reportatives only indicate that the speaker was informed about the situation described in the sentence by someone else, but the speaker uses their own words to do so. If quotatives do not fit in with reportatives, then where do they go in the table in (8)? I would like to propose that quotatives differ from reportatives in terms of the proximal-distal dimension. That is, quotatives are a special instance of speaker-external access to a situation that can be viewed as proximal in two ways. The Speaker not only indicates that they were present at a speech act by someone else, but they also present the quoted speech act in the (proximal) common ground between speaker and hearer. By contrast, using a hearsay/reportative marker would indicate that the situation described in the sentence is situated outside of the proximal space between speaker and hearer: the user of a reportative marker was neither present at the situation described, nor was there necessarily a speech act involved. This analysis places quotatives and witness markers in the same box in the table: sensory evidentials only indicate that the Speaker was present at the situation (sensory/witness), while quotatives indicate more specifically that the Speaker was present at a speech act that they reproduce in their own speech act.

The observation that quotative and reportative differ along the proximal-distal dimension can be most easily seen in direct and indirect speech in English. In (10a), direct speech can only be referred to by the proximal demonstrative

9 The evidential status of reportatives and quotatives has been extensively discussed by e.g. Faller 2002, Anderbois 2014, Korotkova 2017.
this, but not by the distal demonstrative that. By contrast, the complementizer for indirect speech in (10b) is based on the demonstrative distal that, rather than on proximal this.

(10)   a. Sue said (this/ *that): “It is raining.”
   b. Sue said (that/ *this) it is raining.

The relation between reportative and quotative is however not always defined along the proximal-distal dimension. In some languages, like Cuzco Quechua, there are markers that do double duty as quotatives and reportatives. Korotkova (2017) proposes to treat these as homophonous markers, but that seems unfortunate, as it treats the relation between reportative and quotative as entirely accidental. Under the analysis presented here, it is sufficient to say that the Quechua reportative/quotative marker –si is underspecified for the proximal-distal distinction.

10 Conclusion

Summing up, I have called into question the status of evidentiality as an autonomous grammatical category. I propose a programmatic alternative in which evidentiality arises as a result of the recombination of abstract properties of other, more basic grammatical categories. Since such recombinations often carry the ‘baggage’ of the original grammatical categories, a certain overlap and vagueness in evidential meanings is to be expected, in addition to the core set of evidential meanings expressed in Table 1.

Speas (2010) was the first to propose that evidential meanings can be reduced to more abstract primitives of the grammar, even though she still grant evidentiality the status of a separate grammatical category. I have tried to complement and constrain Speas’s (2010) analysis by a ‘spatial’ analysis of evidentiality that appeals to the dimensions of proximal/distal and Speaker-internal/ Speaker external.

The ‘recycling’ perspective on evidentiality should be seen as an application of Occam’s razor: an analysis that manages to derive evidentiality with a more parsimonious set of elementary categories should be viewed as superior to an analysis that needlessly adds taxonomic distinctions.

More broadly, a ‘recycling’ perspective on evidentiality could provide us with information on the limits and organization of the functional domain, a program
akin to that of Wiltschko (2014). If evidentiality recycles independently motivated, more basic categories of human language, such a result would illuminate our understanding of the elementary categories of language as grammatical tools for representing the origin and reliability of our statements. Such a program would allow us to distinguish the elementary categories of grammar from categories that are derived from these. Moreover, it would illustrate the surprisingly versatile uses to which these elementary categories can be put. This approach would afford a new window into the way in which the language faculty accommodates the diversity of the world’s languages in terms of a limited set of abstract elementary categories that can be put to a variety of uses.

References


An emerging veridical complementizer *to čto* in Russian

Alexandra Simonenko & Anastasiia Voznesenskaia

1 A new complementizer

This squib explores an innovation in the complementizer system of Modern Russian: a form *to čto*. The new form is illustrated in (1), which contrasts with the standard *čto* in (2) by the presence of the *to* element.

![Image](image-url)

(1) mne govor'-at, *to čto* eso zakrytyj jamajskij kvartal v centre
Pharaoh in an interview to Yurij Dud', 5:34

(2) esli im nрав-it-s'a, *čto* nравyt's'a
Pharaoh in an interview to Yurij Dud', 31:06

Complementizer *to čto* is string identical to a construction involving a demonstrative introducing a subordinate clause. Specifically, the new complementizer is homographous with a combination of a demonstrative pronoun *to “that”* and a standard complementizer *čto “what”*, as in (3).

(3) Menya udiv-ilo *(to)*, [čto ey] bylo vsego dva god-a

![Image](image-url)

3This is a small token of deep gratitude for the inspiration Liliane gives on the academic, professional, personal, aesthetic, and joie de vivre level.

https://www.youtube.com/watch?v=VXK12UaxtSM

3This also may be precisely the construction that gave rise to the new complementizer.

4The use of a demonstrative with free relatives in Russian is contingent on contrastive inter-
Similar constructions seem to be found in German and English, as (4) illustrates.

(4) Wir bedauern es, dass die Schwimmer nicht bereit sind, die konzeptionellen Dinge mitzutragen.

“We regret it that the swimmers are not willing to share the conceptual issues.”

In this squib we summarize what is already known about the new phenomenon and make some novel observations based on the analysis of recent interviews and speaker judgements. We also sketch a way to approach its analysis, arguing that the new complementizer signals the truth of the embedded proposition with respect to the beliefs of one of the speech act participants or of a third-party.

2 Not a demonstrative to followed by a complementizer čto

The distribution of to čto as a complementizer only begins to be studied as the innovation seems to have emerged during the last decade. In the literature, the claim that we are witnessing the emergence of a new complementizer was first made by Korotaev (2016), who observed that to and čto in this case form an international unity and that to čto can combine with a wider class of matrix predicates than a demonstrative followed by čto.

As noticed by Korotaev (2016), another hallmark of to čto that sets it apart from demonstratives with subordinate clauses is the insensitivity of to in to čto to the selectional properties of the main predicate. While in (5), which features a new complementizer, the form to stays invariant, in (6) to is a demonstrative obligatorily introduced by a preposition and taking a prepositional case, as required for non-propositional arguments by the predicate uveren (“sure”).

(5) ya uveren to čto Timur rukovodstvov-al-sya tvorčesk-imí

I sure to čto Timur direct-PST-REFL creative-INSTR.PL

pretation: for a felicitous use of a demonstrative, the proposition denoted by the free relative should be part of a set of contextually relevant alternatives (Kobozeva 2013). For instance, a natural continuation for (5) can be the following: (I was surprised that she was only two years old) not that she could sing so well.
zadač-ami
goal-INSTR.PL
“I am sure that Timur was motivated by creative goals.” Yuriy Dud’ in an interview with BadComedian, 28:19

(6) Ya uveren v tom, čto Timur rukovodstvov-al-sya tvorčesk-imi
I sure in that.PREP COMP Timur direct-PST-REFL creative-INSTR
zadač-ami.
goal-INSTR.PL
“I am sure that Timur was motivated by creative goals.”

We notice that the two constructions contrast with respect to a number of other properties as well. In particular, unlike the new complementizer, demonstratives introducing subordinate clauses are syntactically mobile and can occur before the main predicate, as (7) shows.

(7) (To,) [čto ey bylo vsego dva god-a], menya udiv-ilo
that COMP she.DAT was only two year-GEN.PL I.ACC surprise-PST.3SG
“I was surprised that she was only two years old.” lit. “That she was only two years old surprised me.”

In contrast, the complementizer in question, just as the standard complementizer čto, has to be postverbal. Consider the ungrammatical (8) (cf. (1)).

(8) *to čto eto zakrytyj jamajskij kvartal v centre mne govor'-at
that to čto this closed Jamaican neighbourhood in centre I.DAT say-PRS.3PL
Intended: “they tell me that this is a closed Jamaican neighbourhood in the center”

Another striking illustration of the complementizer function of to čto are cases where it appears in a clause introduced by an unrelated demonstrative pronoun, as in (9).

(9) vse uveren-y v tom, to čto my pobed-im?
everybody sure-PL in that.PREP to čto we win-FUT.1PL
“Is everyone sure that we are going to win?” From “Vozrast nesoglasiya” episode 4, 10:19

---
5 https://www.youtube.com/watch?v=bag-O-KrswA
6 We are very thankful to Anton Simonenko for his help with grammaticality judgements.
7 https://www.youtube.com/watch?v=emDTVDn0e8M
3 Not an equivalent to the complementizer čto either

Although we fully agree with Korotaev (2016) that to čto should be analyzed as a complementizer, we argue that this is not a full parallel to the standard čto. First, unlike čto, to čto can be used to embed partial questions, as in (10), where to čto combines with počemu “why”.

(10) i vot on vyš-el i mne govor-yat to čto počemu
and here it come.out-PST.3SG and I.DAT say-PRS.3PL to čto why
oni pohoži?
they similar
“and here it [the album] comes out and they tell me why are they simi-
lar?”
Pharaoh in an interview to Yuriy Dud’, 25:30

This is completely out for the standard čto, (11).

(11) i vot on vyš-el i mne govor-yat, (*čto) počemu
and here it come.out-PST.3SG and I.DAT say-PRS.3PL COMP why
oni pohoži?
they similar
“and here it [the album] comes out and they tell me why are they simi-
lar?”

Second, not all matrix predicates that embed clauses headed by the standard čto, are compatible with to čto. For instance, to čto is unacceptable with mečtat’ (“to dream”), (12), pridumat’ “make up, invent”, (13), spasibo “thank you (that)”, (14), orat’ (“to yell”), (15), razočarovan (“to be disappointed”), (16).

(12) ya konečno mečt-al, (*to) čto mama kogda-nibud’ menya
I of.course dream-PST to čto mom some.time I.ACC
voz’m-et, otved- et v futbol’n-uyu škol-u
take-FUT.3SG lead-FUT.3SG in football-PREP school-PREP
“I of course dreamt that my mom will take me to a football school”. Yuriy Dud’ in “Večerniy Urgant”, 3:48

8https://www.youtube.com/watch?v=VBXGvCvECIg. The grammaticality judgement for to čto is added by our consultant.

265
(13) Ja pridum-al (*to) čto u mašin-y bud-et gibridnaya korobka.
I invent-PST to čto at car-ACC will-3SG hybrid transmission
“I made up that the car will have a hybrid transmission.”

(14) spasibo tebe bol’šoe (*to) čto ty ne vzj-al menya v sportivnyj
thank you big to čto you not take-PST l.ACC in sport
smoking.pause
“A big thanks to you that you did not hire me for the “Sport break”.
Dud’ in “Večerniy Urgant”, 14:41

(15) ona ora-la (*to) čto pol toľko čto pokrasi-l-i.
She yell-past to čto floor just paint-PST-PL
“She yelled that the floor has just been painted.”

(16) ya razočarovan (*to) čto lekciju otmeni-l-i
I disappointed to čto class cancel-PST-PL
“I am disappointed that the class has been cancelled.”

(17) Danil požalova-l-sja (*to) čto ya pro nego zab-yl
Danil complain-PST-REFL to čto I about him forget-PST
“Danil complained that I forgot about him.”

The distribution of the two complementizers with respect to a sample of matrix predicates is summarized in table 1.

9https://www.youtube.com/watch?v=VBXGvCvECIg. The grammaticality judgement for
to čto is added by my consultant.
A relative veridicality marker

We make the following descriptive generalizations about the use of to čto based on table 1. First, to čto is not used with predicates which do not attribute the embedded proposition to anyone’s beliefs. Such are orat’ (“to yell”) and spasibo (“thank you that”). Second, it is not used with predicates of creation (of propositional content or possible worlds), such as mečtat’ (“to dream”) and pridumat’ (“to invent”). Third, it is out with emotive predicates which appear to trigger a presupposition that the proposition denoted by the subordinate clause is part of the common ground and asserts an emotional attitude towards that proposition, such as razočarovan “to be disappointed” and požalovat’sya “to complain”.

In view of these generalizations, we suggest that čto signals veridicality, where veridicality is understood as in (18).

(18) A propositional operator F is veridical iff from the truth of Fp we can infer that p is true according to some individual x (i.e. in some individual x’s epistemic model).

Specifically, we propose that to čto signals that the proposition denoted by the complement clause is entailed by the belief worlds of a contextually determined belief source (not necessarily identical to the agent of the main predicate). This means that its use requires that it be contextually possible that there is an epis-
temic source relative to the proposition in question. This condition is violated by predicates of mental creation, in which case propositions are “generated” by an epistemic source rather than being entailed by belief worlds. This also explains the incompatibility of to čto with strongly factive emotive predicates ra-
zočarovan “to be disappointed” and požalovat’sya “to complain”. Factivity, un-
derstood as veridicality relative to the belief’s of the speaker and the hearer at the same time, is arguably more relevant than veridicality relative to a given epis-
temic source. Therefore marking the latter with factive predicates is infelicitous, as its use may trigger an implicature that factivity does not hold.

In future work we hope to link these observation with the structural and se-
mantic typology of complementizer types (Baunaz 2018) as well as to position to čto in a bigger picture of Russian complementizers (Hansen et al. 2016).

References

Hansen, B., Letuchiya, A. & Błaszczyk, I. 2016. Complementizers in Slavonic (Russian, Polish, and Bulgarian). Complementizer semantics in European languages Berlin: Mou-
ton de Gruyter 175–223.
Kobozeva, I.M. 2013. Conditions on the use of a correlative pronoun “to” before com-
plement clauses with “chto”-complementizer [uslovja upotrebljenija “to” pered prida-
Korotaev, N.A. 2016. Soyuz to čto v ustnoy russkoy reči. Russian State University for the Humanities.
Clause-initial subject doubling in Dutch dialects
(Or: Liliane was right after all)

Jeroen van Craenenbroeck & Marjo van Koppen

Abstract

This paper revisits the debate between van Craenenbroeck & van Koppen (2002, 2007) on the one hand and Haegeman (1990, 1992, 2004) on the other regarding the status of clause-initial subject doubling in Dutch dialects. Contrary to our earlier work, we conclude that Haegeman (2004) is right in drawing a principled distinction when it comes to this type of doubling between the dialect of Lapscheure and that of Wambeek. Stronger still, we argue that Haegeman (1990, 1992)'s original position, whereby the first subject element in clause-initial subject doubling is a clitic (rather than a weak pronoun) is indeed the correct analysis. After presenting some arguments—both old and new—in favor of this position, we conclude by considering the broader theoretical implications of this analysis.

1 Introduction: the bone of contention

This paper revolves around a particular configuration of pronominal subject doubling which is attested in various Dutch dialects. It is exemplified in (1) and (2).

(1) Ze guı̂ zaai.
   she.DEFICIENT goes she.STRONG
   ‘She’s going.’

Wambeek Dutch

(2) Ze goa zie.
   she.DEFICIENT goes she.STRONG
   ‘She’s going.’

Lapscheure Dutch

In both these examples the subject pronoun is expressed twice, but interpreted only once. As such, they represent cases of pronominal subject doubling. Two further characteristics will play a central role in the remainder of this paper. First, one instantiation of the subject is a strong pronoun (zaai and zie), while the other is deficient (ze). Second, these examples display what one could call clause-initial subject doubling, in that the first subject pronoun is also the first element of the clause. Subject doubling is by no means restricted to this sentence type, though: both in the dialect of Wambeek and in that of Lapscheure, subject doubling...
doubling also occurs in embedded clauses (the a-examples in (3) and (4)) and inverted main clauses (the b-examples in (3) and (4)).

(3) a. da se zaai guit. 
   that she.DEFICIENT she.STRONG goes 
   ‘that she’s going.’ 
   b. Gui se zaai? 
   goes she.DEFICIENT she.STRONG 
   ‘Is she going?’

   Wambeek Dutch

(4) a. da se zie goat. 
   that she.DEFICIENT she.STRONG goes 
   ‘that she’s going.’ 
   b. Goa se zie? 
   goes she.DEFICIENT she.STRONG 
   ‘Is she going?’

   Lapscheure Dutch

One of the main bones of contention between Haegeman (2004) (henceforth H) and van Craenenbroeck & van Koppen (2007) (henceforth VC&VK) concerns the question of whether (1)–(2) and (3)–(4) represent a unified phenomenon or not. VC&VK claim that they don’t:

while (3)–(4) are representative of a phenomenon they call clitic doubling, the doubling in (1)–(2) is of a fundamentally different nature, which they call topic doubling. H, on the other hand, contends that for the Lapscheure data, this split is uncalled for and hence that (2) and (4) represent the same type of doubling, in particular clitic doubling. Closely related to this analytical difference is the question of the pronominal status of the first subject element in (1)–(2). VC&VK argue that while clitic doubling (the examples in (3)–(4)) always involves the combination of a clitic and a strong pronoun (in the sense of Cardinaletti & Starke (1999)), topic doubling never contains a clitic. This means that the deficient pronoun at the beginning of the clause in (1)–(2) is a weak pronoun. Haegeman (1990, 1992), on the other hand, claims that the deficient pronoun in all the examples in (2) and (4) are subject clitics. The only difference is that while in (4) the clitic cliticizes onto C in narrow syntax, in (2) it does so at PF. A central argument in this debate comes from examples like (5) (van Craenenbroeck & van Koppen 2007:157).

(5) a. { We / * Me } komme waailie mergen.
   we.WEAK we.CLITIC come we.STRONG tomorrow
   ‘We are coming tomorrow.’

b. da { * we / me } waailie mergen kommen.
   that we.WEAK we.CLITIC we.STRONG tomorrow come
   ‘that we are coming tomorrow.’

The reason for this difference is the Verb Second requirement, which regulates that specCP should be filled in a declarative main clause in Dutch (dialects), so the clitic can only move to C at PF, after it has satisfied V2 at Spell-Out. Haegeman (2004:132–133) reinterprets this difference as a difference in pronominal status: ze is a clitic in (4) but a weak pronoun in (2). As van Craenenbroeck & van Koppen (2007:157) point out, however, this claim risks undermining the unified account of clitic doubling in (2) and (4). Moreover, given that we will argue that ze is in fact a clitic in both (2) and (4), we will stick with Haegeman (1990, 1992)’s original analysis in the remainder of the paper and not follow Haegeman (2004)’s reinterpretation.

270
In the first person plural the dialect of Wambeek makes a morphological distinction between its clitic and its weak pronoun, and as the examples in (5) make clear, it is the clitic pronoun that shows up in embedded clauses and inverted main clauses (clitic doubling contexts according to VC&VK), whereas the weak pronoun appears in subject-initial main clauses (topic doubling). According to VC&VK, exactly the same distribution underlies the examples in (1)–(4), save for the fact that the deficient form ze is homophonous between a clitic and a weak pronoun. In a sense, then, those examples are less informative than the ones in (5). H, on the other hand, contends that within Lapscheure (where no minimal pairs like the ones in (5) can be constructed) there is no evidence in favor of a different status of ze and so the conclusion drawn on the Wambeek data does not carry over to Lapscheure.

In this paper we revisit this debate, and conclude, contrary to our earlier work, that Haegeman (1990, 1992, 2004) is right, and that the type of doubling illustrated in (2) is of a fundamentally different nature than the one in (1): while the latter is a case of topic doubling, the former represents clitic doubling, exactly like the examples in (3)–(4). In support of this position, we present four arguments, which we lay out in detail in the next section.

2 Arguments for a clitic doubling analysis

2.1 Introduction

In this section we present four arguments in favor of the hypothesis outlined in the previous section. Some of these arguments are known from the literature (though not always explicitly acknowledged as such), others are new.

2.2 Coordination

As pointed out by (Cardinaletti & Starke 1999:169), weak pronouns differ from clitics in that the former can but the latter cannot be a shared subject of a predicate coordination. If the two dialects under consideration here differ in the pronominal status of clause-initial deficient pronouns, we expect that difference to be manifested there as well. As the examples in (6) (van Craenenbroeck & van Koppen 2000:12) and (7) show, this expectation is borne out.

(6) *'k Spelen op de piano en zingen altijd geweldig.
    I.DEFICIENT play on the piano and sing always great
    INTENDED: 'I play the piano and always sing wonderfully.' Lapscheure Dutch

Fortunately, we have not always been wrong-headed: in van Craenenbroeck & van Koppen (2000:40) we conclude, partly based on data from the Lapscheure dialect, "that West Flemish does not have [topic] doubling". We should have stuck to that initial position.
While in Lapscheure Dutch the deficient pronoun 'k cannot occur as the shared subject of a predicate coordination, its identically sounding counterpart in the Wambeek dialect can. This suggests that there is a difference in pronominal status between these two elements: 'k is a clitic in (6), but a weak pronoun in (7). This is consistent with a clitic doubling analysis of (2) and a topic doubling analysis of (3).

2.3 Geographical distribution

A second way to distinguish clitic doubling from topic doubling concerns their geographical distribution. As discussed in detail by de Vogelaer & Devos (2008), clitic doubling and topic doubling are typically found in different geographical areas. First consider the map they provide of clitic doubling:

![Geographical distribution of clitic doubling](image)

Figure 1: Geographical distribution of clitic doubling (de Vogelaer & Devos 2008: 256)

There is a core clitic doubling area, which consists of the provinces of French Flanders, West Flanders, and East Flanders. In addition, there appear to be remnants of a clitic doubling system, where we find what looks like first and second person doubled pronouns. Following Pauwels (1958), Nuyts (1990), De Schutter (1994) and de Vogelaer (2005), we assume that these are not actual cases of (clitic) doubling, but that they involve reanalysis of originally clitic doubled forms as non-doubled, positionally restricted strong pronouns. As such, we will not discuss these forms any further in the rest of the paper.

The distribution of topic doubling can be represented as follows:

Topic doubling is concentrated in the provinces of Flemish Brabant and Antwerp, with extensions into the northeast of East Flanders. When comparing Figure 1 with Figure 2...
it becomes clear that the geographical distribution of clitic doubling and topic doubling is quite distinct—nearly complementary even. This means that we can use the distribution of the phenomena to shed light on the status of the examples in (1)–(2): if they pattern as in Figure 1, we are led to an analysis in terms of clitic doubling, while if they show the distribution in Figure 2, we are dealing with topic doubling. Interestingly, de Vogelaer & Devos (2008) do precisely this. Their findings, as well as their interpretation of these findings, can be found in Figure 3.
The combination of a deficient pronoun followed by the finite verb followed by a strong pronoun occurs in a geographical area that includes both the clitic doubling area in Figure 1 and the topic doubling area in Figure 2. When faced with this distribution, de Vogelaer & Devos (2008:262) conclude that “a non-uniform analysis may actually provide a better understanding of the data”. In other words, the pattern ‘deficient pronoun + finite verb + strong pronoun’ should be analyzed as clitic doubling in the area marked in black in Figure 3 and as topic doubling in the area marked in grey. Given that Lapscheure belongs to the black area and Wambeek to the grey one, these findings support the hypothesis that the examples in (1)–(2) should receive a non-uniform analysis.

2.4 Tripling

If the Lapscheure data in (2) and (4) both represent clitic doubling, then this dialect has only one mechanism for doubling a subject.

Wambeek Dutch, on the other hand, has two such mechanisms: topic doubling in (1) and clitic doubling in (3). This opens up the possibility of these two mechanisms co-occurring in a single sentence. Specifically, while the clause-initial subject is doubled by a strong pronoun (topic doubling), that strong pronoun could in turn be doubled by a clitic (clitic doubling), thus leading to a three-fold instantiation of the subject, or tripling. As pointed out by Haegeman (1992:66) and shown in (8), tripling is disallowed in the dialect of Lapscheure. In Wambeek Dutch, on the other hand, tripling is fine (see (9)).

(8) *Ze goa zie.
    she.DEFICIENT goes she.CLITIC she.STRONG
    INTENDED: ‘She’s going.’  Lapscheure

(9) Ze gui zaai.
    she.DEFICIENT goes she.CLITIC she.STRONG
    ‘She’s going.’  Wambeek

The contrast between these two examples suggests that while Wambeek Dutch has two doubling mechanisms at its disposal and hence allows for their co-occurrence, Lapscheure Dutch only has one. In van Craenenbroeck & van Koppen (2006) we explore this contrast for a slightly larger number of dialects and arrive at the same conclusion. The dialects under investigation there are represented in Figure 4.

The group of dialects uninspiredly referred to as “A-dialects” in Figure 4 are like Lapscheure Dutch in that they disallow pronominal tripling, while the so-called “B-dialects” are like that of Wambeek in allowing the subject to be doubled twice within one sentence. Like the map in Figure 3, then, these facts suggest that the contrast between Lapscheure and Wambeek under investigation in this paper is part of a larger generalization, whereby (roughly) the provinces of French Flanders, West Flanders, and East Flanders behave like Lapscheure Dutch in disallowing topic doubling, while Flemish Brabant and Antwerp are like Wambeek Dutch in having this construction.

4We’re abstracting away from so-called topic marking here, as this is not really a doubling phenomenon. See van Craenenbroeck & Haegeman (2007), de Vogelaer & Devos (2008), and Haegeman (2008) for discussion.
2.5 Meaning differences

Van Craenenbroeck & van Koppen (2002) coined the term ‘topic doubling’ for the example in (1) (and—wrongly—also for the one in (2)) to highlight the fact that the first subject element occupies a position in the CP-domain, a region typically associated with information-structural effects. One of the examples they present in support of this position is the minimal pair in (10) (van Craenenbroeck & van Koppen 2002:295).

(10) a. Een vrouw komt a kaffee binn.  
   ‘A woman enters a bar.’
   a woman comes a bar in
b. Een vrouw komt zaai a kaffee binn.  
   ‘Women usually enter a bar.’
   a woman comes she.STRONG a bar in
   #’A woman enters a bar.’
   Wambeek Dutch

In (10a) the indefinite subject is not doubled and a non-specific indefinite reading is possible, while in (10b) this reading is absent and only a generic interpretation is available. This shows that topic doubled subjects behave like topicalized constituents: both are incompatible with a non-specific indefinite reading.

If the Lapscheure example in (2) represents a case of clitic doubling, not topic doubling, we expect this type of doubling not to be sensitive to information-structural considerations. As far as we know, this has not been looked into in any detail for the Lapscheure dialect (though see Haegeman (1996:335n2) and (Haegeman 2004:127) for occasional remarks about the meaning of this type of doubling), but Uittenhove (2015) examines pre-
cisely this issue for the closely related dialect of Bredene. Through an extensive questionnaire, he examines whether information-structure has an effect on the use of subject doubling in the dialect of Bredene. He distinguishes between four contexts—new information focus, givenness topic, contrastive focus, and contrastive topic—and examines to what extent these contexts favor subject doubling. The main finding of his research is quite unequivocal: “De zwak-sterkverdubbeling (...) wordt in alle contexten aanvaard. Zowel als focus als als topic krijgt de zwak-sterkverdubbeling van het subject hoge scores” (Uittenhove 2015:68). This is exactly what we would expect in the context of this paper: if West Flemish lacks topic doubling altogether, then there should be no interaction between subject doubling—i.e. clitic doubling—and the information-structural properties of the subject.

2.6 Conclusion

We have just reviewed four pieces of evidence that all point to the same conclusion: the data in (1) and (2) should not be given a unified account. More specifically, while in Wambeek Dutch topic doubling as in (3) contrasts with the clitic doubling examples in (3), the Lapscheure data in (2) and (4) seem to uniformly point towards a clitic doubling analysis. In addition, we have also shown that there are reasons to think that this interdialectal difference is symptomatic of a larger split, which (roughly) contrasts the Flemish dialects with the Brabantic ones.

3 Conclusion and broader implications

The one-sentence summary of this paper is clear and easy: we were wrong and Liliane was right in the analysis of the Lapscheure data in (2). More generally, it looks like topic doubling, while a real phenomenon, does not extend all the way into the Flemish dialects, but is stopped in its tracks at the Flemish-Brabantic border. The broader implications of this regional divide are, we believe, well worth exploring (see van Craenenbroeck & van Koppen (2016) for relevant discussion).

Another consequence of the proposal developed here relates to Haegeman (1990)’s original analysis of subject doubling in Lapscheure Dutch. Recall from footnote 2 above that the analysis of an example like (2) differed from those in (4) in that the clitic only cliticized to C at PF. The reason for this difference was the V2-requirement of Lapscheure Dutch: the preverbal position has to be filled by an XP at the point of spell-out. However, if we are right that the preverbal subject element in (2) is a true clitic, then the preverbal position is occupied by a head, not an XP in this example. The key to understanding this conclusion, we believe, lies in Jouitteau (2010)’s reclassification of V2-, SVO- and VSO-languages into one single category of X(P)-VSO. The V2-constraint—however implemented and probably more aptly called the X(P)-VSO-constraint—prohibits the finite verb

---

5 Both the Lapscheure dialect and the Bredene one are part of so-called coastal West Flemish, see Taeldeman (2013).

6 “Deficient-strong subject doubling is accepted in all contexts. Both as focus and as topic this type of doubling receives high acceptability ratings.”
from being leftmost in the clausal phase (Jouitteau, 2011:10), and one of the ways in which this constraint can be respected is by merging a head to the left of the finite verb. Lapscheure Dutch can do this in a context like (2), but Wambeek Dutch cannot. While this distinction might seem far-fetched or arbitrary at first, van Craenenbroeck (2011) argues that there is another context in which the clause-initial position is occupied by a head in Lapscheure Dutch, namely in expletive constructions (see Haegeman (1986), Grange & Haegeman (1989) for extensive discussion):

(11) T ĉȜĕđ zyn gisteren drie studenten gekomen.

‘Three students came yesterday.’ Lapscheure Dutch

Van Craenenbroeck (2011) argues that the expletive element t should not be analyzed as a reduced form of the third person neuter personal pronoun het ‘it’, but rather as a West Flemish analogue to Welsh or Breton clause-initial particles, with which they share many morphosyntactic properties (Jouitteau 2008, 2010, 2011, Borsley et al. 2007, Willis 1998, 2007). Viewed from this perspective, both the expletive example in (11) and the clitic doubling example in (2) are representative of the X-VSO-nature of Lapscheure Dutch. Working out this connection in more detail is a topic for further research, though.

References


Handout for an invited talk at the 26th Comparative Germanic Syntax Workshop.


Here, there, and (every)where

Guido Vanden Wyngaerd

1 Introduction

The words here, there, and (every)where (henceforth HTW) are traditionally taken to be adverbs. Evidence discussed in Burton-Roberts (1991), however, shows that they behave distributionally like PPs. For reasons of space, I do not recapitulate this evidence here, but summarise it schematically in (1):

(1) |   | Adv | PP | HTW |
---|---|---|---|---|
complement of V | X | ✓ | ✓ |
modify Adj/Adv | ✓ | X | X |
postmodify N | X | ✓ | ✓ |
complement of P | X | ✓ | ✓ |
take PP complement | X | ✓ | ✓ |
take right, straight, just | X | ✓ | ✓ |
locative inversion | X | ✓ | ✓ |

Burton-Roberts (1991: 171) concludes from these data that HTW are prepositions. In terms of their meaning, however, HTW do not just correspond to simple prepositions, but are richer semantically. This is seen in the analysis of Katz & Postal (1964), who argue that HTW derive from an underlying PP-like structure:

(2) here : at this place
    there : at that place
    where : at what place

3See also Kayne (2009), who argues that here and there are licensed in a structure with silent nouns (THIS here PLACE, THAT there PLACE). The list in (2) may be extended with complex expressions like everywhere and somewhere, as well as other expressions of direction and/or location, like back, away, home, upstairs, downstairs, abroad, aboard, apart, aside, together, which behave like HTW. I shall not further discuss those here, however.
Aarts (2013) takes HTW to be PPs, which is more in line with their semantics. The conclusion that HTW are PPs is still too general, however: HTW correspond with a subclass of the PPs, namely those with a locative or directional meaning. The locative meaning in particular appears from the fact that some of the environments in (2) require a locative PP, such as the complement of go or put, or the phenomenon of locative inversion.

The conclusion that what look like single words in fact correspond to complex structures, viz. PPs of the type in (2), provides an interesting case for the nanosyntactic mechanism of phrasal spellout, by which a syntactically complex constituent can be spelled out by a single lexical item. At the same time, there is also some evidence for decomposing HTW. Again looking at both the forms and the meanings in (2), we see that there is also a deictic (this/that) or a wh-element (what) present. This element is responsible for the semantic differences between here (deictic proximate), there (deictic distal), and where (interrogative, or indefinite in complex forms like somewhere, everywhere; see Haida 2007). The forms also suggest a decomposition of the type {h/th/wh}-ere, with on the one hand an exponent h-/th-/wh-, which is responsible for the deictic/wh part of their meaning, and an exponent -ere on the other, which spells out a different set of features relating to location and direction. The spatial triplet h-ere/th-ere/wh-ere has a temporal cognate in the triplet now/th-en/wh-en, which (except for now) is subject to a similar decomposition, with the second part (-en) referring to time rather than to location/direction. In this short paper, I focus on the second part of HTW (-ere), which for convenience I continue to refer to as HTW. I argue that -ere realises the locative/directional part of the meaning of here, there, and where. I defer a further discussion of h-/th-/wh- to another place and time.

2 Movement, Direction and Location

Before proceeding with the analysis of HTW, I need to discuss prepositional expressions of location and direction. Certain types of P only have a locative meaning (e.g. in), while others are directional (e.g. to). The difference shows up most clearly in stative predications, such as postnominally or with be (Déchaine et al. 1995):

\[(3) \quad \begin{array}{l}
a. \quad \text{a train in}_{\text{loc/to}}_{\text{dir}} \text{ Paris} \\
b. \quad \text{This train is in}_{\text{loc/to}}_{\text{dir}} \text{ Paris.}
\end{array}\]

French à ‘at’ only has a locative sense:

Taking Dir and Loc to stand in a containment relation (with Dir containing Loc), the difference between locative in and directional to can be conceived of as a difference in size, as shown in (5):

(5)  
<table>
<thead>
<tr>
<th></th>
<th>Dir</th>
<th>Loc</th>
<th>Place</th>
</tr>
</thead>
<tbody>
<tr>
<td>in</td>
<td>Paris</td>
<td></td>
<td></td>
</tr>
<tr>
<td>to</td>
<td>Paris</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This approach allows an account of the otherwise surprising fact that, combined with certain types of motion verbs, both English in and French à seem to express direction:

(6)  
a. She went/came/fell/jumped/flew in the water.  
b. Ce train va à Paris.  
   this train goes at Paris  
   ‘This train goes to Paris.’

This fact can be explained by assuming that the Dir element can be realised by a motion verb (Fábregas 2007, Caha 2010), allowing a locative preposition to spell-out the remainder of the functional sequence. This is shown schematically in (7):

(7)  
<table>
<thead>
<tr>
<th></th>
<th>Dir</th>
<th>Loc</th>
<th>Place</th>
</tr>
</thead>
<tbody>
<tr>
<td>be</td>
<td>in</td>
<td>Paris</td>
<td></td>
</tr>
<tr>
<td>go</td>
<td>in</td>
<td>the water</td>
<td></td>
</tr>
<tr>
<td>aller</td>
<td>à</td>
<td>Paris</td>
<td></td>
</tr>
</tbody>
</table>

Not all motion verbs are able to spell out Dir: English walk, run or dance, when combined with in, only have a locative, and no directional, sense, because neither the verb nor in can realise Dir. A directional preposition like to is needed, as shown in (8) and (9).

(8) She walked/ran in inLoc/to inDir the park.
Following Levin et al. (2009), I call verbs like *go/fall/jump* verbs of directed motion (or directional verbs for short), and verbs like *walk/run/dance* manner of motion verbs (see also Talmy 1975, 1985 on path-framed vs satellite-framed languages).

3 Analysis

With this in mind, let us return to HTW. This section aims at demonstrating that English HTW can occur in all the slots where locative and directional PPs can occur. This distributional pattern is accounted for by assuming that HTW are the phrasal spellout of a constituent corresponding to a locative/directional PP.

The locative sense of HTW appears in stative predications like (10a) and (10b), and it also appears in sentences where HTW combine with directional verbs, like (10c). Under the analysis developed in the previous section, *DIR* is realised by the verb in such a case. As a result, HTW shrink to realising only *LOC*, in spite of the directional sense expressed by (10c). This is shown schematically in (11).

(10)    a. The pharmacy is there.
 b. They live here.
 c. She came here first and then went there.

(11)    | DIR | LOC | PLACE |
        | be   | in   | Paris |
        | be   | there|       |
        | go   | there|       |

The directional sense of HTW becomes apparent from (12), with manner of motion verbs.

(12)    She walked/ran there_{LOC/DIR}.

This sentence is in fact ambiguous, in a manner which is reminiscent of (8) above: there, with *DIR* left unexpressed and *LOC* expressed by *in*, the sense was locative, but with *DIR*+*LOC* realised by the preposition *to*, the meaning was directional. In this case, the ambiguity of (12) suggests that there is a syncretism between directional and locative HTW. In line with our earlier assumptions, we assume...
that directional HTW spell out a larger constituent than locative HTW, as shown in (13):

(13)          | DIR | LOC | PLACE |
            | walk | there |  
            | walk | there |  

Standard Superset Principle logic accounts for this syncretism: the lexical entry for HTW contains the features $\text{DIR}+\text{LOC}+\text{PLACE}$, and given that there is no competing lexical item that just spells out $\text{LOC}+\text{PLACE}$, the larger lexical item may spell out the smaller syntactic structure that is contained in its lexical entry. Finally, observe that in addition to the features $\text{DIR}$ and $\text{LOC}$, HTW also spell out an abstract noun meaning $\text{PLACE}$. In this respect, HTW differ from now/then/when, which realise an abstract noun TIME.

4 Conclusion

I have argued that HTW are not adverbs, nor prepositions, nor PPs, but are decomposable into a deictic/wh part $h-/\text{th-}/\text{wh}-$ and a locative/directional part -ere. The -ere part is the phrasal spellout of an abstract set of features expressing direction and location, and the abstract noun $\text{PLACE}$. How the two parts of HTW and their corresponding feature sets connect with each other is a matter which I hope to return to in future work.

References


Grammaticalisation processes in Flemish Sign Language

Mieke Van Herreweghe

Following Hopper & Traugott (2003 [1993]: 232), grammaticalisation can be defined as “the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions.” Grammaticalisation processes have not been studied very extensively in sign languages yet. Pfau & Steinbach (2006) give a very interesting survey of studies that have focused on grammaticalisation processes in sign languages, but Flemish Sign Language (VGT) was not one of them. Within the Deaf community in Flanders about 5000 - 6000 people (Loots et al. 2003) claim to have Flemish Sign Language as their first or principal language. After lengthy negotiations, VGT was officially recognized by the Flemish Parliament in April 2006. VGT clearly is a fully-fledged sign language in its own right, and is genealogically related to amongst others French-Belgian Sign Language (LSBF), French Sign Language (LSF), American Sign Language (ASL) and Sign Language of the Netherlands (NGT). The common ancestor of these daughter sign languages is Old French Sign Language (OFSL). However, it is impossible to use historical data to look at grammaticalisation paths since there simply are very few historical grammatical data as OFSL was never written down. Consequently, the method to be used is that of internal reconstruction which is a procedure for inferring part of the history of a language from material available for a synchronic description of the language on the basis of paradigmatic allomorphy.

1 Grammaticalisation clines

For spoken languages, grammaticalisation processes have been described along a number of structural changes or clines. The following examples will show that these clines can be found in VGT as well.
### 1.1 Morpho-syntactic fusion

According to Bybee (1985) morphosyntactic fusion treats the relationship between syntax (or sentence structure) and morphology (or internal word structure) or to put it in Givón’s (1971: 413) terms “Today’s morphology is yesterday’s syntax”. Pfau & Steinbach (2006: 87) state that “sign languages only have very few (if any) instances of type 2-grammaticalization (i.e. from free to bound grammatical morpheme)”. I would like to argue that at least one example can be found in VGT (and probably also in other sign languages) with respect to negative affixation. Clearly, certain negative verb signs in VGT have developed from a combination of a positive sign and the negative adverb NOT (see Figure 1) resulting in a positive verb stem followed by a negative affix which consists of a twisting movement of the wrist.

![Figure 1: Negative adverb NOT (with a left to right horizontal sweeping movement) (picture taken from Van Herreweghe & Vermeerbergen 2006: 245)](http://gebaren.ugent.be/alfabet.php?id=23011)

Examples are for instance the verb signs BELIEVE-NOT, WANT-NOT (see Figure 9), CAN-NOT (see Figure 10) and the deverbal adjective sign UNKNOWN which is the same as the noun STRANGER (see Figure 2).

![Figure 2: UNKNOWN or STRANGER (http://gebaren.ugent.be/alfabet.php?id=23011)](http://gebaren.ugent.be/alfabet.php?id=23011)

---

**Decategorialisation**

Decategorialisation refers to the evolution of open class lexemes in a primary or major category to closed class lexemes in a secondary or minor category. As has been described for other sign languages VGT also has instances of the evolution of the gesture for “strong” being lexicalised into the (ad)nominal sign STRONG/POWER (with an upward movement) and then grammaticalised into the modal verb CAN (with a downward movement).

![Figure 3. STRONG/POWER (http://gebaren.ugent.be/alfabet.php?id=18674)](http://gebaren.ugent.be/alfabet.php?id=18674)

![Figure 4. CAN (http://gebaren.ugent.be/alfabet.php?id=22028)](http://gebaren.ugent.be/alfabet.php?id=22028)

Since a similar path has been described for ASL (Janzen and Shaffer 2002) and other OFSL related sign languages (Wilcox, 2004) the assumption can be that at least the lexicalisation but maybe also the grammaticalisation already took place in OFSL. Another example of decategorialisation in VGT is the evolution of the adjectival/adverbial sign READY into an aspectual marker READY, quite similar to what has been described for FINISH in ASL (Janzen 1995).

![Figure 5. READY (http://gebaren.ugent.be/alfabet.php?id=22239)](http://gebaren.ugent.be/alfabet.php?id=22239)

Furthermore it would appear that the subordinating conjunction BECAUSE has developed from the nominal sign REASON, which may be similar to its NGT counterpart (Pfau & Steinbach 2007: 40).
1.2 Decategorialisation

Decategorialisation refers to the evolution of open class lexemes in a primary or major category to closed class lexemes in a secondary or minor category. As has been described for other sign languages VGT also has instances of the evolution of the gesture for “strong” being lexicalised into the (ad)nominal sign STRONG/POWER (with an upward movement) and then grammaticalised into the modal verb CAN (with a downward movement).

Figure 3: STRONG/POWER (http://gebaren.ugent.be/alfabet.php?id=18674)

Figure 4: CAN (http://gebaren.ugent.be/alfabet.php?id=22028)

Since a similar path has been described for ASL (Janzen & Shaffer 2002) and other OFSL related sign languages (Wilcox 2004) the assumption can be that at least the lexicalisation but maybe also the grammaticalisation already took place in OFSL. Another example of decategorialisation in VGT is the evolution of the adjectival/adverbial sign READY into an aspectual marker READY, quite similar to what has been described for FINISH in ASL (Janzen 1995). Furthermore it would appear that the subordinating conjunction BECAUSE has developed from the nominal sign REASON, which may be similar to its NGT counterpart (Pfau & Steinbach 2006: 40).

However, in VGT the following paths can be discerned which have not (yet) been described for other sign languages:
- lexical verb GIVE → light verb GIVE → auxiliary/preposition GIVE functioning as recipient marker (see below)
- nominal sign EXAMPLE → conjunction introducing a conditional clause
Decategorialisation refers to the evolution of open class lexemes in a primary or major category to closed class lexemes in a secondary or minor category. As has been described for other sign languages VGT also has instances of the evolution of the gesture for “strong” being lexicalised into the (ad)nominal sign STRONG/POWER (with an upward movement) and then grammaticalised into the modal verb CAN (with a downward movement).

Since a similar path has been described for ASL (Janzen and Shaffer 2002) and other OFSL related sign languages (Wilcox, 2004) the assumption can be that at least the lexicalisation but maybe also the grammaticalisation already took place in OFSL.

Another example of decategorialisation in VGT is the evolution of the adjectival/adverbial sign READY into an aspectual marker READY, quite similar to what has been described for FINISH in ASL (Janzen 1995).

Furthermore it would appear that the subordinating conjunction BECAUSE has developed from the nominal sign REASON, which may be similar to its NGT counterpart (Pfau & Steinbach 2007:40).

1.3 Phonological reduction

The grammaticalised element is frequently phonologically reduced compared to its non-grammaticalised counterpart. For VGT we can find phonological reduction in for instance the negative affixation mentioned above. Another example is the sign GIVE which in its full lexical verb form is signed with a horizontal movement from the agent to the recipient. It is usually (but not compulsory) formed with spatial agreement, i.e. the movement of the sign starts at the locus of the agent and ends at the locus of the recipient.

In its reduced form, when used as a preposition functioning as recipient marker (see below), it is formed with a short horizontal forward wrist-flipping movement starting from the signer without any spatial agreement.
1.4 Reduction of syntactic freedom

Heine et al. (1991) and Lehmann (1985) claim that grammaticalisation entails a limitation of syntactic freedom since a lexical element can be moved around more freely while a grammaticalised element is more limited in its syntagmatic relation to other sentence elements. In VGT GIVE as an auxiliary or preposition is again a good example of this structural change since both are always positioned right in front of the indirect object or recipient. This is quite striking since for VGT it is frequently very difficult to detect clear word order rules, since there are mostly only word order tendencies (Vermeerbergen, 2004).

2 Principle of divergence

Following the principle of divergence (Hopper, 1991: 24) it is possible that the different forms that can be found on a grammaticalisation cline exist next to each other at the same time while the variants can be put in a hierarchy from less to more grammaticalised. One such example is the variety with respect to the verb sign GIVE (itself a lexicalisation of a classifier construction). The examples can...
all be found in Devriendt (2009) where they are discussed more elaborately.

1. Classifier construction or incorporated classifier: SOMEONE BOOK GIVE-classifier for book TO BOY. In this construction a classifier handshape or book is incorporated in the verb sign GIVE.

2. Conventionalised sign GIVE: SOMEONE BOOK GIVE\textsubscript{3l} TO BOY. Here the citation form of the verb sign GIVE is a conventionalized form although it is possible (but not compulsory) that spatial agreement is applied.

3. GIVE in a verb sandwich construction (Fischer & Janis 1990): WOMAN GIVE PRESENT \textsubscript{1} GIVE-classifier for present \textsubscript{3r}. In this example the first GIVE is the conventionalised citation form without any spatial agreement and the second one has an incorporated classifier handshape for present and there is spatial agreement.

4. Light verb (cf. Butt 2004) GIVE: GIRL GIVE BOY STROKE\textsubscript{self} STROKE\textsubscript{3}. Here again the citation form of GIVE is used, without any spatial agreement and with a certain level of semantic bleaching since a stroke cannot be handed over from one person to another.

5. Auxiliary GIVE: INDEX\textsubscript{m} RABBIT \textsubscript{1} GIVE\textsubscript{3l} MAN SHOOT\textsubscript{ml}. In this example (in which the rabbit is shooting the man) again the citation form of GIVE (with spatial agreement with the locus of the recipient) is used followed by the recipient. The label auxiliary is used here since its use seems to be very similar to the NGT auxiliary ACT-ON (Bos 1994).

6. Preposition GIVE: RABBIT GIVE MAN SHOOT\textsubscript{ml}. The label preposition is given here since contrary to the auxiliary GIVE there is no spatial agreement and there is a strong phonological reduction (as described above).

3 Unidirectionality?

Most researchers would claim that grammaticalisation paths are unidirectional, i.e. that there is development from a full lexical element into a functional grammatical element. It appears that at least some counterexamples to this general rule can be found in VGT. For a more detailed description of these signs we refer to Huy\textsubscript{3} (2008).

There is a possible development of the negative modal auxiliary WANT-NOT into a full lexical verb “cannot be bothered”. Both signs are formed in exactly the same way, but the non-manual part of the sign differs because the latter has a compulsary mouth gesture (i.e. not referring to any Dutch word) while the former can be accompanied by the mouthing “wil-niet” (want-not).
Another example could be the possible development of the negative modal auxiliary CANNOT plus MORE into a full lexical verb meaning “cannot take it anymore”. Again the latter has to be formed with a compulsory mouth gesture while the former can be accompanied by the mouthings “kan niet meer” (can not more).

Some researchers would regard these counterexamples as a challenge to the principle of unidirectionality, while others would take a more careful stand.

In conclusion, Pfau & Steinbach (2006: 87) state “that sign languages employ exactly the same grammaticalization paths as do spoken languages. That is, the pathways proposed in the literature are modality-independent”. From the small-scale study on VGT which has been reported on here, it seems to be able to corroborate this for VGT as well.
References


“Stemmen uit het verleden”
Het transcriptieproject bij de Vakgroep Taalkunde—Nederlands van de Gentse Universiteit

Jacques Van Keymeulen & Anne-Sophie Ghyselen

1 Inleiding

Prof. Lilianne Haegeman is een internationaal gereputeerde syntactica, die meermaals het West-Vlaamse dialect van het dorp Lapscheure, haar moedertaal, als onderzoeksobject gekozen heeft. Syntactisch onderzoek van de oraal overgeleverde traditionele taalvariëteiten is echter niet makkelijk, mede doordat syntactische constructies via opvragingen lastig te eliciteren zijn. Syntactische constructies zijn, in tegenstelling tot bijvoorbeeld lexicale items, weinig saillant en bovendien zijn ze dikwijls optioneel en geconditioneerd door situationele context. In de huidige tijd kan men bij opvragingen ook de invloed van de standaardtaal hoe langer hoe minder uitsluiten. Nagenoeg elke veldwerker heeft ondervonden dat een respondent bij een expliciete vraag een bepaalde constructie als ongrammaticaal kan beoordelen en die dan – op een minder bewaakt ogenblik – toch zelf gebruikt. De Syntactische Atlas van de Nederlandse Dialecten (SAND) legt getuigeis af van een complexe opvragingsmethode in vier rondes, met een inventarisatieronde, schriftelijke en mondelinge vragenlijsten, gecomplementeerd met telefonische navraag (zie Barbiers 2005). Blijkbaar moest een en ander dikwijls gecontroleerd worden.

De bovengemelde problematiek is natuurlijk al lang geleden ingezien. De professoren W. Pée (1903–1986) en V.F. Vanacker (1921–1999), die – als student! – de eerste syntaxis voor een Vlaams dialect schreef (namelijk dat van Aalst; Vanacker s.d.), vatten begin jaren 60 van de vorige eeuw het plan op om voor elke gemeente in Nederlandstalig België (+ Zeeuws- en Frans-Vlaanderen)

2 De verzameling dialectbanden aan de Vakgroep Taalkunde—Nederlands van de UGent


Het idee was om via een zgn. “vrij gesprek” voor elke gemeente ongeveer 45 minuten spontaan gesproken dialect te verzamelen bij een goede dialectspreker. Die dialectspreker moest voldoen aan een aantal objectieve criteria: hoge ouderdom, lage geletterdheid en honkvastheid. Daardoor kwam men in de praktijk vaak bij de boerenbevolking terecht. Elke bandopneming wordt begeleid door een fiche (in A4-formaat), met de biografische metadata van de spreker. Er werd dikwijls gebruik gemaakt van studenten om geschikte informatie in te vullen op de fiche.  


Het is jammer dat die metadata niet ingesproken waren op de band zelf (die bevat enkel plaats en datum van opname), zodat men telkens de fiche erbij moet nemen om de opname goed te kunnen evalueren.
machten te zoeken – meestal werden mannen uitgekozen. De opnames gebeurden bij de zegspersoon thuis, dikwijls met een dialectsprekende tussenpersoon – in die tijd konden de meeste studenten nog vrij goed een lokaal dialect spreken.

De collectie bevat momenteel 783 banden voor 550 plaatsen, vooral uit het Vlaamse dialectgebied (= Frans-, West-, Oost- en Zeeuws-Vlaanderen). Naar het oosten toe (provincies Antwerpen, Vlaams-Brabant en Limburg) neemt het aantal geluidsopnames af - uit die provincies komen gemiddeld minder studenten naar de Gentse universiteit. Voor sommige dorpen zijn er verschillende opnamen, vooral voor die die aan de basis van een syntactische licentiaatsverhandeling hebben gelegen. De meeste opnames zijn gemaakt bij mensen die tussen 1885 en 1910 geboren zijn; de oudste spreker (uit Bossuit) is geboren in 1871. In veel gevallen werd een zegspersoon uitgekozen die een zeldzaam traditioneel ambacht had uitgeoefend. Voor Frans-Vlaanderen werd een speciale inspanning gedaan: daar heeft men een opname gemaakt voor elke gemeente waar nog inheemse Vlaamssprekenden aangetroffen konden worden.

De verzameling had in de eerste plaats een taalkundig oogmerk: hoewel op de begeleidende fiche plaats was ingeruimd voor notities omtrent “Onderwerp van het gesprek”, werd de inhoud van het gesprek niet of maar heel gebrekkig aangeduid. Toch is het zo dat het geluidscorpus de grootste verzameling levensverhalen uitmaakt van zeer laag geschoolde mensen, die niet alleen het begin van de ‘moderne tijd’ hebben meegemaakt (aanleg van elektriciteit, eerste fiets, eerste auto ...), maar ook twee wereldoorlogen hebben doorstaan. Vele verhalen leggen getuigenis af van de grote armoede en onwetendheid waarmee een groot deel van de Vlaamse plattelandsbevolking zo’n honderd jaar geleden heeft moeten leven.

3Meestal werd met een Telefunken M 25 of met een Revox-toestel opgenomen (BASF-banden, type LGS 35, op spoelen van 18 cm diameter met snelheid 19 = 48 minuten geluid). De sectie Nederlandse Taalkunde heeft nog steeds een gebruiksklaar Revox-toestel waarmee de originele banden beluisterd kunnen worden.


5Het is erg jammer dat men voor de verzameling niet met de KULeuven heeft kunnen samenwerken.

6Zie bijvoorbeeld Van Keymeulen (1975).

De digitalisering hield volgende zaken in. Allereerst werden de 783 geluidsbanden gedigitaliseerd. Een technisch medewerker van de toenmalige vakgroep Nederlandse Taalkunde, Rieke Willems, heeft ongeveer 500 banden op zeer professionele manier naar wav-bestanden omgezet. Toen mevr. Willems wegens een reorganisatie naar een andere sectie van de vakgroep werd gemuteerd, is de rest van de banden gedigitaliseerd door een gespecialiseerde firma.

Die subsidies werden verworven via de organisatie “Variaties vzw. Koepelorganisatie voor Dialecten en Oraal Erfgoed in Vlaanderen".
Op de tweede plaats werden alle metadata in een database ondergebracht. Op de derde plaats werd door vrijwilligers van de meeste banden korte inhouden gemaakt, voorafgegaan door trefwoorden, zodat de verzameling inhoudelijk ontsloten werd. Momenteel zijn er ongeveer 700 korte inhouden in de database ingebracht. Op de vierde plaats werden alle transcripties (typoscripten) – voor zover aanwezig – ingescand en in de database ingevoerd; het zijn er 318 in totaal. Die transcripties zijn niet fonetisch, maar ‘woordelijk’, d.w.z. dat de dialectische uitspraak in een vernederlandste vorm is weergegeven. De typoscripten werden in de jaren 60 en 70 gemaakt door jobstudenten aan de hand van een aantal richtlijnen – die met wisselende ijver en kunde werden opgevolgd. De verzameling originele banden is bij het universiteitsarchief ondergebracht; kopieën van de digitale versie ervan (+ kopieën van de fiches met metadata) zijn gedeponeerd zowel bij de Koninklijke Academie voor Nederlandse Taal- en Letterkunde (KANTL) in Gent als bij het Meertens Instituut in Amsterdam.²

⁸Voor het verkrijgen van een kopie van een opname moet contact opgenomen worden met de sectie Nederlandse Taalkunde van de UGent. Elke gebruiker moet dan een gebruiksovereenkomst ondertekenen waarin hij de aansprakelijkheid opneemt voor het vrijwaren van de privacy van de zegslieden.
4 Het transcriptieproject

Naarmate de traditionele dialecten verdwijnen, verhoogt de waarde van de verzameling geluidsoptnames. De specifieke waarde ervan heeft vooral te maken met de spontaneïteit van de gesprekken, hoewel ook daar met de ‘observer’s paradox’ rekening gehouden moet worden. In de huidige digitale tijden, heeft men echter nood aan een ander type transcripties dan een (beperkte) aantal typoscripten, die niet digitaal doorzoekbaar zijn.

In het academiejaar 2016-2017 werd daarom aan de sectie Nederlandse Taalkunde voor de bachelorscripties een onderzoekslijn opgestart met als bedoeling digitaal doorzoekbare transcripties te maken, die gealigneerd bleven met het geluid. De bedoeling is ervaring op te doen en een transcriptieprotocol te verfijnen dat dan in de toekomst gebruikt kan worden bij wetenschappelijke projecten. Die ervaring heeft in elk geval al geleerd dat de transcriptie van de verzameling erg dringend is: jonge mensen hebben al heel wat moeite om het ouderwetse dialect te begrijpen en ook hebben ze geen voeling meer met de gespreksonderwerpen, zoals bijvoorbeeld de manier waarop vlas werd geoogst, of hoe eertijds palingen werden gevangen. Ook wordt getracht de verzameling uit te breiden – vooral naar het oosten toe: de studenten werden gevraagd een bandopneming te maken voor een plaats waarvoor er nog geen bestond. Het is immers nog steeds mogelijk honkvaste mensen te vinden met een lage opleiding die talig gesocialiseerd zijn vóór de jaren 60 van de vorige eeuw, toen door de veralgemeende toename van de mobiliteit, de verhoging van de scholingsgraad en de introductie van de massamedia een massale taalverandering is op gang gekomen. De zegspersoon moet nu dus vóór 1940 geboren zijn, zodat de kans vrij hoog is dat de zegspersoon nog een traditioneel dialect als primaire moedertaal heeft verworven.

De studenten maken voor de transcripties gebruik van het programma ELAN en transcriberen het dialect op twee niveaus: licht vernederlands en zwaar vernederlands. Met lichte vernederlandziging wordt bedoeld dat enkel de fonologie naar het Nederlands wordt aangepast, d.w.z. dat de morfologie en syntaxis van het dialect ongewijzigd blijft, maar dat alle woorden worden neergeschreven alsof ze in het Nederlands bestonden. Clitische vormen worden aan het grondwo-

---

9De observer’s paradox valt voor de banden goed mee omdat de meeste zegspersoons werklijk niets anders dan dialect konden spreken.
10In het academiejaar 2016-2017 en 2017-2018 gingen er telkens een viertal studenten aan de slag.
11Een zegspersoon is momenteel idealiter dus ongeveer 80 jaar of ouder.
ord geschreven, maar gescheiden door een #. De transcriptie met zware vernederlandsing is op die lichte vernederlanding gebaseerd. Hier worden de functiewoorden – ook die in de clitische vormen – vertaald naar het Nederlands. Die zware vernederlanding is nodig om achteraf softwaretools te kunnen gebruiken voor (semi-automatische) Parts of Speech-tagging, lemmatisering, en andere vormen van annotatie. Het transcriptieprotocol voor een dergelijke manier van werken is vrij complex; de studenten hadden gemiddeld ongeveer een uur tijd nodig om 5 minuten geluid te transcriberen.

Om het onderscheid tussen de lichte en zware vernederlanding te illustreren, geven we hieronder een stuk uit een transcriptie uit Ieper weer. In de opname, die dateert uit 1967, vertelt een toen 62-jarige vrouw (die we het label N72.sp1 geven) over de oorlog, haar carrière en pensioen, het leven in leper en de jeugd van tegenwoordig. Het excerpt is kort, maar getuigt meteen van een aantal interessante syntactische dialectverschijnselen, zoals dubbele negaties met het partikel en, variaties op het introducerende er en subjectverdubbelingen.

(1) N72.sp1 ja a#je#t nu ingaat
N72.sp1.ndl ja als je het nu ingaat
N72.sp1 wuk zit er nu veel in de jongheid?
N72.sp1.ndl wat zit er nu veel in de jongheid?
N72.sp1 zeg mijn dat.
N72.sp1.ndl zeg mij dat.
Int_VV ze zijn allemaal...
N72.sp1 allemale nie nee#s. allemale nie wi. ke zeggen dat op
N72.sp1 ndl allemaal niet nee ze. allemaal niet wi. ik zeg dat op
jou nie wi jongen.
jou niet wi jongen.
Int_V ja ja m...
N72.sp1 maa algelijk vele.
N72.sp1 ndl maar algelijk veel.
Int_VV t#is toch veel veranderd.
N72.sp1 je zout#ier moeten zijn in de zomer eni meneer.
N72.sp1 ndl je zou hier moeten zijn in de zomer eni meneer.
N72.sp1 op de Remparden.
N72.sp1 ndl op de Remparten.
Int_VV ja.
N72.sp1 ze komen van t#schole met under carnassière. ze
N72.sp1 ndl ze komen van de school met hun carnassière. ze
staan en likken en frikkelen.
staan en likken en frikkelen.
N72_sp1 en ze moeten nie peinzen dat ze nunder generen wi.
N72_sp1_ndl en ze moeten niet peinzen dat ze zich generen wi.
N72_sp1 ze generen under nie wi.
N72_sp1_ndl ze generen zich niet wi.
N72_sp1 en sommigste mensen en durven me nunder
N72_sp1_ndl en sommige mensen en durven met hun
kinders... kleine kinders op de vestingen nie meer gaan.
kinderen... kleine kinderen op de vestingen niet meer gaan.
Int_VV ja ja.
N72_sp1 voilà.
N72_sp1_ndl voilà.
N72_sp1 da... da betaamt algelijk nie.
N72_sp1_ndl dat... dat betaamt algelijk niet.
Int_VV ja ja betaamt nie...
N72_sp1 da#ze doen da ze willen. ze zijn jong.
N72_sp1_ndl dat ze doen dat ze willen. ze zijn jong.
N72_sp1 maa da#ze#t doen waa da ze nie ezien en zijn.
N72_sp1_ndl maar dat ze het doen waar dat ze niet gezien en zijn.
Int_VV voilà. andere mensen...
N72_sp1 m#en wijder ook jong eweest.
N72_sp1_ndl we hebben wij ook jong geweest.
N72_sp1 m#en wijder ook evrijd.
N72_sp1_ndl we hebben wij ook gevrijd.
N72_sp1 is#t geen waar?
N72_sp1_ndl is het geen waar?
N72_sp1 maa me deden#t algelijk nie in de presentie van
N72_sp1_ndl maar we deden het algelijk niet in de presentie van
alleman.
alleman.
Int_VV ja ja ja.
Int_VV ja ja da#s just.
N72_sp1 maa nu... ze zijn nie meer beschaamd voor nie...
N72_sp1_ndl maar nu... ze zijn niet meer beschaamd voor nie...
N72_sp1 tot op strate.
N72_sp1_ndl tot op straat.

303
Toekomstige projecten?

De onderzoekslijn voor de bachelorscripties heeft als bedoeling wetenschappelijke projecten voor te bereiden. Gelukkig zijn er nu in principe ook wetenschappelijke fondsen beschikbaar om databases aan te leggen voor toekomstig onderzoek. Daarom is het zaak om transcripties te maken en publiek beschikbaar te stellen die een zo ruim mogelijk onderzoeksveld kunnen bedienen. De transcriptie van de volledige verzameling van 780 banden – of een representatieve selectie ervan – zou in elk geval een taalmonument tot stand brengen waarmee onverschillig welke onderzoeker van dialecten of historisch Nederlands zijn voordeel zou kunnen doen.

Prof. Anne Breitbarth (sectie Duitse taalkunde UGent) is erin geslaagd FWO-financiering te verwerven voor een pilootproject met de transcriptie van 30 opnames (+ annotatie daarvan), netjes verdeeld over de verschillende dialectgroepen in Nederlandstalig België en Frans- en Zeeuws-Vlaanderen. Ook hopen we het beste van een andere aanvraag voor de transcriptie van alle Frans-Vlaamse opnames. De bewerking van Frans-Vlaanderen is uiteraard het dringendst; het Vlaams in Frankrijk is zeer op de terugweg; de oude dialecten zijn hoe langer hoe moeilijker te begrijpen. Er is ook aan andere universiteiten (o.a. Leuven en Antwerpen) grote belangstelling voor de Gentse verzameling. Het is de bedoeling binnen afzienbare tijd een zgn. Herculesproject in te dienen bij
het Fonds voor Wetenschappelijk Onderzoek – Vlaanderen om de verzameling – meer dan een halve eeuw jaar na het ontstaan ervan – definitief te ontsluiten.

References


305
Diary null subjects: an analogy with imperatives?

Andrew Weir

Liliane and I first met at the LAGB meeting in 2008. I wasn’t presenting; I was there to meet her. Caroline Heycock had sent Liliane my undergraduate dissertation, about null subjects in spoken and written English, which had naturally drawn heavily on Liliane’s work on diary drop (Haegeman 1990, 1997, 2007, 2013, 2018, Haegeman & Ihsane 1999, 2001); and we talked for some time about my dissertation and about the possibility of my coming to Ghent for doctoral study. Well, I took a detour via Massachusetts to get to Ghent, but I got there eventually; and can honestly say that my year there with Liliane as my supervisor was one of the most enjoyable I have had as a linguist. Liliane was and is a fantastic mentor, co-author, and friend. However, when I was in Ghent, we got distracted by the properties of response particles in West Flemish, and never actually returned in a systematic way (though we’ve exchanged a lot of emails about it) to the phenomenon that ‘got us together’ in the first place – diary drop. This squib is an attempt to remedy that – and an invitation to Liliane for us, as we have thought we should for a while, to join forces and solve this knotty problem for good.

We start with an empirical question: are there person restrictions on null subjects in ‘reduced written register’ (RWR)? First-person null subjects are indubitably OK, and third-person null subjects also are given an appropriate discourse (cf. e.g. Haegeman 2007: 96), but this is somewhat difficult to determine for second person; diary context does not lend itself to second-person subjects, and interrogatives (which might more naturally have second-person subjects) are independently ruled out with null subjects (in any person) in English diaries (Haegeman 1997 a.o.). Berthelot (2017: 88) states that second-person singular null subjects are ruled out in diary contexts in French, but in Weir (2018: 159), I suggested that null second-person pronouns were grammatical in English RWR, on the basis of the below example.
(1) (comments on a student’s essay)
\[ \emptyset_{[2p]} \] Don’t need to go into so much detail here.

But I should have been more careful. The problem is that (1) is grammatical in spoken register too, something which is the result of a quite different process, a prosodically-governed operation that can target a wide range of deaccented utterance-initial material (Napoli 1982, Weir 2012). That raises the possibility that (1) is simply a rendering in text of that spoken string. To check if diary drop per se can delete second-person pronouns, one needs to check examples like (2), which are ungrammatical in speech (see Weir (2012) for why). And interestingly, in my judgment, they are also degraded in written register (e.g. text messaging), even if the ‘topic’ of the utterance is plausibly the addressee. For me, the examples in (2) contrast fairly sharply with first-person dropped subjects (3).

(2) You shouldn’t have done that. ??\[0\]_{[2p]} Are going to regret it. / ??\[0\]_{[2p]} Will regret it. / ??\[0\]_{[2p]} Have been reported to the police.

(3) \[0\]_{[1p]} Am going to regret this. / \[0\]_{[1p]} Will regret this. / \[0\]_{[1p]} Have been reported to the police.

If this is right – and I should caution here that I have not done any corpus work to back up my intuitions – then it seems that second-person subjects can not undergo diary drop. This distinction between first- and second-person does not immediately fall out from extant analyses, as far as I can see. (I’ll return to third person briefly below.) I want to sketch here an alternative line of attack. Consider the below properties of diary drop (from Haegeman 2007: 102):

(4) Diary null subjects are almost always root subjects: they are marginal in embedded contexts without a complementizer, and ruled out under complementizers or in adjunct or relative clauses.
   a. \[0\] Am going to the gym later.
   b. ??\[0\] Think \[0\] will go to the gym later.
   c. *\[0\] Think that \[0\] will go to the gym later.
   d. *\[0\] Was sore after \[0\] had gone to the gym.

\(^1\) I am assuming here that the grammatical properties of e.g. text message register are fundamentally the same, at least when it comes to subject drop, as diary register. This might not be a safe assumption.

\(^2\) Modulo the more liberal, ‘Bridget Jones’ dialect described in Haegeman & Ihsane (1999, 2001). The properties of that dialect remain to be accounted for. The marginality (as opposed to complete ungrammaticality) of (4b) represents my judgment.
Diary null subjects are strongly ungrammatical in construction with subject-auxiliary inversion or with wh-fronting.

a. *Am ∅ just going mad?
b. *What was ∅ thinking?
c. *Only then will ∅ do that.

Diary null subjects are degraded in construction with argument topicalization – although (in my judgment) the degradation is not as strong as with SAI or wh-fronting.

a. ??More problems, ∅ don’t need. (Thrasher 1977: 83)

Diary null subjects are however permitted in construction with left-peripheral adjuncts, or with left-dislocation.

a. Tomorrow ∅ will go to gym.
b. If you are hungry, ∅ have left some biscuits in the cupboard.

A fact that I don’t believe has previously been noted is that this constellation of facts lines up nearly perfectly with constraints on (English) imperatives.

Only permitted in root contexts (marginally in certain complementizerless embedded clauses)

a. Fix the problem.
b. ?He said fix the problem. (Crnič & Trinh 2009)
c. *He said that fix the problem.
d. *I demand that fix the problem.

No wh-fronting/auxiliary fronting

a. Don’t touch! – Don’t touch what?/*What don’t touch?

b. Never eat this./*Never do eat this. (cp. negative inversion/do-support in declaratives: Never do I eat this, Henry 1995: 68f.)

---

3 Not constraints on imperative null subjects – the grammaticality of the examples in (9)-(11) does not change if the subject is overt – but constraints on the syntax of imperatives generally.

4 One might wonder if don’t has in fact undergone subject-aux inversion here (with the unpronounced subject), but Zhang (1991), Henry (1995), Rupp (2007), Weir (2013) give reasons to believe that imperative don’t is base-generated above the subject.
Argument topicalization marginal (but better than wh-movement/SAI)

a. ??These instructions, read carefully.
b. ??These cookies, don’t touch.

Adjunct topics and left-dislocation OK

a. Tomorrow, go to the gym.
b. If you’re hungry, eat some biscuits.
c. These cookies, don’t touch ’em.

We might hope to capitalize on these similarities, and on the fact that imperatives, exceptionally among clause types in English, allow null (second-person) subjects. One common way of understanding, or encoding, this latter fact is to propose that some functional head is – in imperatives only – endowed with second-person φ-features (see recently e.g. Jensen (2003), Bennis (2006), Zanuttini (2008)). These features can enter into some (obligatorily very local) relationship (such as Agree) with a pro subject with the same features, licensing its non-pronunciation. Such a representation ‘encodes the addressee in the syntax’, as Zanuttini (2008) puts it. Authors differ in where they put the φ-features, but it seems plausible to me that a construction restricted to imperatives should have its etiology localized to the head that encodes the force of the clause (cf. Bennis 2006: sec. 3.1), so this is where I propose the φ-features are in (12).

(12) \[ \text{ForceP Force}_{\text{imp}},[2p] \quad \text{[vP pro}_{[2p]} \quad \text{[vP eat your dinner]]} \]

Suppose, now, that just as imperatives encode the addressee in the syntax, written-register declaratives encode (or can encode) the speaker. That is, these registers allow a declarative Force head to (optionally) be endowed with first-person φ-features. This Force head can enter into the same (obligatorily local) relation with a pro subject and license its non-pronunciation.

(13) \[ \text{ForceP Force}_{\text{decl}],[1p]} \quad \text{[TP pro}_{[1p]} \quad \text{[TP am going to the gym later]]} \]

One might speculate that these φ-endowed Force heads are highly restricted in their distribution (only to root or perhaps certain ‘embedded root’ environments,

\footnote{Though the data are variable here. Potsdam (1998: ex. 82) states that sentences like The tie, give to Bob; the aftershave, give to Don! are fully grammatical. They do sound considerably better than those in (10), for reasons I don’t understand.}

\footnote{It needs to be worked out how this aligns with proposals that both speaker and addressee are always encoded in the left-peripheral layer, e.g. Speas & Tenny (2003), Sigurðsson (2004), Sigurðsson & Maling (2010), Haegeman & Hill (2013).}
cf. Portner et al. (2014) on embedded imperatives); that the relevant relationship between the Force head and the pro subject would be disrupted, in both imperatives and diary sentences, in cases where material has moved to the left periphery; and that the obligatory localness of the relation between Force and pro rules out e.g. null objects. Together, then, this is what leads to the correlation between diary null subjects and imperatives shown in (4)–(11). One might also speculate that the diary context provides a ‘natural’ environment for something like this, a declarative Force head endowed with [1p] features, to develop; diaries are prototypically sequences of declaratives where the speaker (writer) is recounting their actions. Once this head was independently available in speakers’ lexicons for the diary register, it might also be plausible to assume that it could easily ‘jump’ to other written registers, such as text message register.

All of the preceding paragraph is highly speculative, and much remains to be worked out. One glaring lacuna is the handling of dropped third-person subjects:

(14) a. $\emptyset_{[1p]}$ Saw Bill yesterday. $\emptyset_{[3p]}$ Was absolutely furious about the bill he'd received.
   b. $\emptyset_{[3p,expl]}$ Is raining. (Haegeman & Ihsane 1999: 121)

To explain (14a), we might appeal to Weir (2018)'s proposal for (third-person) object drop, which proposes the existence of a null D in reduced written register, giving rise to completely silent third-person pronouns. In principle this could be independent of the licensing of null first-person pronouns. However, I crafted that system specifically to exclude the possibility of null expletive pronouns (which I argue do not occur in object position in RWR); so (14b) remains an issue. And there are several other places where the basic idea needs to be fleshed out. This squib, then, shouldn’t be taken as anything definitive. It’s rather something on the order of a tentative proposal – and an invitation to Liliane that, if she finds the ideas here intriguing, we might think about developing them together, and solving the problem of diary drop for good, or at least for a while.

If this is right, it would imply that, if a structure includes $\phi$-endowed Force, the relevant relationship between Force and the subject must be established whether or not the subject is pronounced (because the imperative examples in (9)–(11) have the same status whether or not the subject is pronounced). On my reading of Zanuttini (2008) (who puts the action in a Jussive head rather than Force), this is true for the abstraction relationship that Jussive establishes over the subject, but Zanuttini also discusses (in her section 4) cases where she argues that the Agree relationship between Jussive and the subject can be blocked (without leading to ungrammaticality). Working out if what I am saying is compatible with what Zanuttini says is one of many areas where the present proposal will have to be developed.
All the very best, Liliane – hope have enjoyed reading this squib as much as enjoyed writing it!

Referenties


Recording and explaining: exploring the German ditransitive alternation

Klaas Willems, Ludovic De Cuypere & Hilde De Vaere

1. In the introduction to Thinking Syntactically (2006), Liliane Haegeman writes

   that the goal of scientists is not merely taking note of and recording
   certain phenomena and thus ‘knowing’ about them: scientists want
   to explain the phenomena they have observed.

   The argument that in linguistics, too, the ultimate goal is to explain linguistic
   phenomena rather than to merely record them, is particularly pertinent since the
   ‘quantitative turn’ in modern linguistics. In the wake of this turn, many scholars
   have laid great emphasis on amassing data, according to some critics – not only
   generative linguists – to the detriment of ‘real’ explanations. In this discussion
   note, we offer some thoughts on the relation between explanation and the ex-
   tensive recording of data from a ‘moderate’ functional point of view. The paper
   takes the form of a case study in which we consider the variation in form and
   function of sentences with the ditransitive verb geben in present-day standard
   German. This is the subject matter of an ongoing corpus-based research project
   in the General Linguistics section of the Linguistics Department at Ghent Uni-
   versity.

   2. Our starting point is the common assumption that the only ditransitive
   construction in which geben occurs in the standard language is the Indirect Ob-
   ject Construction (henceforth: IOC) with the RECipient coded in the dative and
   the THEME coded in the accusative, e.g.:

   (1) Diese Kommunikation gibt den BürgernREC ein Gefühl der
       SicherheitTHEME.
       ‘This communication gives the citizens a sense of security.’

   All example sentences are drawn from the 42-billion-word corpus DeReKo (Deutsches Refer-

   313
However, corpus research reveals that *geben* also occurs in the Prepositional Object Construction (henceforth POC): while the THEME is still coded in the accusative, the RECIPIENT is headed by the preposition *an*, which in this construction governs the accusative:

(2) Kronauer will eine Liste mit allen Teilnehmern \textsc{theme} \textit{an die Stadt} \textsc{rec} geben.

‘Kronauer wants to give a list of all participants to the city.’

(3) Der Bund muss nicht direkt das Geld \textsc{theme} \textit{an die Familien} \textsc{rec} geben.

‘The federal government does not have to give the money directly to the families.’

This finding contravenes the commonly held view that the IOC/POC alternation does not exist with *geben* in German, whereas the so-called ‘dative alternation’ is common with the corresponding verb in many Germanic languages (cf. Rapaport Hovav & Levin 2008, Haspelmath & Baumann 2013). Note, moreover, that the IOC/POC alternation is well-attested with other ditransitive verbs in German such as übergeben ‘hand over’, zurückgeben ‘give back’, abgeben ‘pass, hand over’, schicken ‘send’, ausleihen ‘lend (out)’, senden ‘send’, übersenden ‘send’, and so on. As a matter of fact, with morphologically complex *geben*-verbs there is no empirical evidence that one variant outnumbers the other.

In the last two decades a great number of formal, functional and cognitive studies have been devoted to the English dative alternation or ‘dative shift’, as it is commonly called in formalist scholarship. The focus has been on such verbs as *give*, *send*, *throw* and *sell*, which in English either occur in the ‘double-object construction’ (*He gave his 24-year-old son an allowance for spending money*) or the ‘to-construction (*He gave the newspaper to his 24-year-old son*). Research into the corresponding alternation in present-day German has been lagging behind (cf. Matzel 1976, Wegener 1985, Proost 2015). In this contribution we discuss some preliminary findings based on ongoing corpus-based research. We limit ourselves to observations on the simplex verb *geben*. Our study is confined to written language (DeReKo).

3. We conducted random searches in DeReKo in order to acquire sufficient data for a comparison of IOC and POC with *geben*, but because POC turned out to be much less frequent than IOC with this particular verb (approximately 4% of the occurrences), we turned to specific queries with the preposition *an* to arrive at a balanced data set. All sample sentences were annotated for a number of factors along the lines of existing corpus studies of the English dative alter-
The factors in our study are: length difference (heaviness) and order of recipient and theme, voice, pronominality, animacy, concreteness, specific verb sense (concrete, propositional, abstract, cf. below), idiomaticity, definiteness, discourse-givenness. The investigation follows up on previous research, which relates the alternation primarily to two well-established types of motivation: universally applicable processing constraints, in particular heaviness considerations (cf. Hawkins 1994), and information structure preferences (Thompson 1995, Bresnan 2007, Bresnan & Ford 2010, Rappaport Hovav & Levin 2008). According to Hawkins’ principle of early immediate constituents, the orders T(HEME)-R(RECIPIENT) and R-T are determined by the relative weight of both objects, the heavier one tending to be placed after the lighter one (one of ‘Behaghel’s Laws’). Thompson’s principle of ‘topicworthiness’ adds to this that more ‘topicworthy’ objects tend to be placed before less ‘topicworthy’ ones, with topicworthiness defined as ‘a cluster of properties’ that influence the packaging of information with regard to the likelihood of a noun phrase being the topic of discussion (Thompson 1995). Thus, pronominal, animate, definite, specific, identifiable, given and short objects are taken to be associated with referents that are more topicworthy than inanimate, indefinite, non-specific, non-identifiable, new and long objects. Based on this topicworthiness principle, IOC with R-T order is expected to occur mainly with topicworthy recipients whereas POC with T-R order mainly with topicworthy themes.

The analysis of N = 1341 sentences with the verb geben reveals that 95.5% of the IOC sentences have R-T order and 4.5% T-R order (712/33) while 99.5% of the POC sentences have T-R order and 0.5% R-T order (594/2). Although German has a relatively free word order compared to English, the correlation between each variant and a specific order of the two objects is nevertheless very similar for give and geben. Logistic regression analysis further shows that the IOC/POC alternation is significantly associated with multiple factors in German (cf. De Vaere et al. 2018). POC is positively associated with recipients that are longer than the theme, collective recipients (e.g., Familie ‘family’, Präsidium ‘executive committee’, Chor ‘choir’) and recipients that can designate both concrete locations and institutions in the abstract (e.g., Polizei ‘police’, Ministerium ‘government department, ministry’) while the themes are generally discourse-given or at least accessible, often pronominal and they tend to designate concrete objects or propositional contents. There is also a significant correlation of POC with passive voice. Conversely, IOC is positively associated with discourse-new and abstract themes and with pronominal, discourse-given and animate recipients that are shorter than the theme. Representative corpus-extracted examples of IOC and POC are

Wir werden unseren Bericht in der ersten Januar-Hälfte an das Präsidium geben. ‘We will deliver our report to the Bureau in the first half of January.’

We also found that one particular use of *geben* in combination with *an* (occasionally also *in, auf oder über*) is strictly confined to POC, viz. when it is used as a phrasal verb in the sense ‘add an ingredient to’ in the context of preparing food, but in this case the prepositional phrase is not a recipient argument; (6) is an example:

Gewürfelte Zwiebeln können roh oder mit heißer Brühe übergossen an den Salat gegeben werden. ‘Diced onions may be added to the salad raw or doused with hot broth.’

Overall, the findings for German *geben* show interesting parallels with the English dative alternation. Predictions in terms of the traditional heaviness considerations and information structure preferences are largely borne out by the *geben* data. Hence, an appropriate quantitative approach, which is able to uncover correlations between various factors, already goes beyond the stage of mere data recording. It supplies us with explanatory clues which are not readily accessible to introspection, bringing to light regularities that would otherwise for a large part remain hidden. Furthermore, given that the research results for *geben* partly match those for *give* (similar results have been obtained for other Germanic languages, e.g. Dutch, Danish and Swedish), the outcome of this part of the study points to fairly strong general tendencies across Germanic languages with regard to the alternation at hand.

4. It would however be premature to conclude that the above quantitative analysis, indispensable though it may be, provides a full explanation of the data. Not surprisingly, a number of functionally oriented approaches to the English dative alternation, which have gained widespread recognition in recent decades, have sought to determine specific semantic and/or pragmatic differences between the two variants. For instance, in Goldberg’s Construction Grammar approach the alternation in English is analyzed in terms of two different argument structure constructions and three different senses. The double-object
construction is assigned the construction meaning ‘successful caused possession’ (X CAUSES Y TO RECEIVE Z, e.g., John gave Mary an apple). It is contrasted with the ‘transfer-caused-motion construction’ (e.g., John gave an apple to Mary) which Goldberg calls a ‘prepositional paraphrase’ of the double-object construction but semantically a metaphorical extension of the ‘caused-motion construction’ (X CAUSES Y TO MOVE Z, e.g., Joe kicked the bottle into the yard) (Goldberg 1995, 2006). An alternative account is proposed in Rappaport Hovav & Levin (2008)’s verb-sensitive approach. They argue that with a verb such as give both the double-object construction and the to-construction convey ‘successful caused possession’ because of the verb’s inherent meaning. By contrast, verbs such as send and throw convey ‘caused possession’ in the double-object construction but ‘caused-motion’ in the to-construction. Moreover, any successful transfer inference is not determined by the meaning of the construction but by the meaning of the verb (e.g., by give and sell but not by send, throw, kick or teach).

It is not possible within the confines of this contribution to expound in detail to what extent these accounts of the English alternation can be applied to German. However, careful analysis of the data shows that this is possible only to a very limited extent. It is imperative that due attention be paid to languagespecific properties of the alternation, for the following reasons. First, German possesses a fully-fledged system of morphological cases, which are found only in remnants in English. Note that the double-object construction also exists in German but it occurs with only a handful of verbs that take two objects in the accusative (e.g., lehren ‘teach’) and the construction is not confined to ditransitive verbs (compare nennen ‘call someone something’). Moreover, in German various prepositions (in, auf, nach, zu, an) are used to designate places in POC, often corresponding to the single preposition to in English. Second, while in English the dative alternation is co-extensive with R-T order and T-R order, this is different in German. Above we pointed out the strong correlation of IOC with R-T order and POC with T-R order in sentences with geben, but it appears that geben is rather exceptional in this respect among German ditransitive verbs. (This is perhaps less surprising in view of the fact that the morphosyntactic behaviour of ‘give’ is notoriously special from a cross-linguistic and typological point of view, cf. Kittilä (2006)). In a random sample of 3353 sentences with 10 alternating ditransitive verbs other than geben, IOC is attested with R-T order in 67% of the cases against 33% for T-R order (1094/533). With POC, T-R order accounts for 96%, R-T order for 4% of the occurrences (1663/63). Thus, IOC regularly occurs with both object orders, unlike POC. Thirdly, while in English not only give-type verbs and send-type verbs but also throw-type verbs (throw, shoot,
*kick, slap etc.*) partake in the alternation, the corresponding verbs in German *(schmeißen, werfen, schießen, treten, stoßen etc.*) occur in POC but not IOC. Only few morphologically complex throw-type verbs such as *zuwerfen* ‘throw at, pass’ and *hinwerfen* ‘throw at’ occur in IOC (cf. Croft et al. 2001). Finally, with English *give*-type verbs the preposition *to* only takes animate complements but not inanimate complements that designate places (*give something to Berlin/the church/the bureau etc.*) are acceptable only metonymically, cf. Rappaport Hovav & Levin 2008. By contrast, with German *geben*-verbs (*geben, übergeben, zurückgeben, abgeben, weitergeben etc.*) *an* can take animate or inanimate complements. It is obvious that these differences between the two languages have to be taken into account when considering functional contrasts between IOC and POC in German.

5. The results of our case study can be summarized as follows. The IOC/POC alternation in German (with whatever verb) is clearly situated at the syntax/semantics/pragmatics interface, with ramifications into language processing. It is therefore necessary, both with regard to the alternating constructions and the verbs that instantiate them, to distinguish 'encoded' meanings from senses that are not encoded but 'inferred', including those senses that obtain by default in normal language use (for the distinction between 'encoded' meanings and 'inferred' senses, including generalized conversational implicatures, cf. Coseriu 1975, Grice 1989, Levinson 2000, among others).

Under this view, a first conclusion is that the IOC/POC alternation in German cannot be explained in terms of a dichotomy between two independent constructions with contrasting encoded meanings. The analysis of the data shows that it is not an encoded semantic property of POC to convey 'caused motion', nor of IOC to convey '(successful) caused possession'. In this respect, the results of our study square with Rappaport Hovav & Levin 2008’s view that the dative alternation in English does not alter the ‘caused possession’ reading if the verb itself, e.g. *give*, lexicalizes ‘caused possession’. However, on the basis of the German data we have to go one decisive step further.

On the one hand, the encoded lexical meaning of *geben* appears underspecified with regard to the three major conventionalized senses (concrete, propositional and abstract transfer, cf. Bresnan et al. 2007) and any specific subsenses that occur in the data (e.g., ‘hand over’ is a subsense of the concrete sense, on a par with ‘administer’, ‘transmit’, and so on). *Geben* does not express concrete transfer in the majority of the occurrences. Propositional and abstract transfer are equally frequent, but the concrete sense is more often attested in POC, as in (7), whereas the abstract sense is favoured in IOC, as in (8), the propositional
sense regularly occurs in both variants, (9).

(7) Sie können die CD dann vervielfältigen und an die Schulen geben. ‘You can then make copies of the CD and give them to the schools.’

(8) Hier wird den Schülern die Gelegenheit gegeben, mit Zeitzeugen zu diskutieren. ‘Here the students are given the opportunity to discuss with eyewitnesses.’

(9) a. Ein Zeuge beobachtete ihn und gab der Polizei Tipps. ‘A witness watched him and gave tips to the police.’
b. Hier konnte Hebisch einen guten Tipp an die Kameraden geben. ‘Here Hebisch was able to give a good tip to the comrades.’

It would be begging the question to construe the uses of geben in the abstract (8) and propositional (9) senses from a putative concrete ‘core sense’ (7). However, ‘caused possession’ is no encoded feature of geben either. A state of ‘possessing something’ is not necessarily intended in ditransitive uses of geben, compare:

(10) Der Kartensitz und Spielverlauf geben dieser Hoffnung keine Chancen. ‘The hand of cards and the course of the game give this hope no chance.’

(11) [Ich] sehe gute Chancen, der Wirtschaft die dringend benötigten Impulse zu geben. ‘I see good opportunities to give much-needed impetus to the economy.’

Rather than invoking figurative extensions on apriori grounds (cf., e.g., Newman (1996)), which have little support in the data, we argue that the underspecified encoded meaning of geben is best paraphrased as a general three-place ‘transfer’ meaning with no specification as to the features ‘possession’, ‘path’, and ‘transfer modality’, in contrast to other core three-place verbs such as schicken, senden, liefern, schenken, and complex geben-verbs.

On the other hand, the corpus investigation also shows that in German, IOC and POC are not exclusively dedicated to their functions in the ditransitive alternation. ‘An AGENT transferring an entity to a RECIPIENT’ is but one event type both variants can designate, which we take as evidence that this three-participant frame is not their encoded meaning but one of their possible senses. Building on typological research of the ditransitive construction (Kittilä 2006, Malchukov et al. 2010, Bickel 2011, Haspelmath 2013 among others), we therefore stress the need to analyze the alternation in terms of a more general configuration (cf. Stefanowitsch 2011) on the systemic level of German grammar. This schematic con-
configuration combines three semantic roles, viz. an AGENT, a THEME, and a GOAL, to a three-argument pattern. IOC and POC have the status of ‘allostructions’ (cf. Cappelle (2006)) that instantiate this configuration rather than being independent constructions in the grammar (unlike the English double-object construction, which has an encoded ditransitive meaning, if previous analyses are correct); nor can one allostruction be considered more ‘basic’ than the other. The two allostructions are moreover partly in complementary distribution in German: give-type verbs and send-type verbs occur both in IOC and POC but throw-type verbs occur in POC (with the exception of a small number of morphologically complex verbs such as zuwerfen, hinwerfen etc.). Furthermore, whereas the GOAL is a spatial goal (DESTINATION) with throw-type verbs, it is either a spatial goal or a possessional goal (RECIPIENT) with give-type verbs and send-type verbs.

Importantly, the occurrences of IOC and POC are no realizations on a one-off basis. Their uses reflect the pervasive role of a handful of conventionalized senses, which in turn correlate with various morphosyntactic, semantic and pragmatic factors as well heaiveness constraints. These correlations concur to establish observable, albeit non-exclusive, tendencies in language use which, although no rules of grammar, are indispensible to understand why on a particular occasion of language use one variant is likely to be chosen whereas the other one is more or less strongly dispreferred. For instance, the tendency for IOC to occur with the abstract transfer sense of geben in combination with an abstract THEME and an animate RECIPIENT contrasts with the no less notable tendency for POC to occur with the concrete or propositional transfer sense of the verb in combination with a RECIPIENT that either designates a collective entity or allows for both a concrete locative or an abstract institutional reading (e.g., Ministerium ‘government department, ministry’).

6. In conclusion, the level of normal language use, situated in between the language system and actual instantiations of language use, is key to a layered account of the alternation (cf. Coene & Willems 2006). This level straddles the stringent dichotomous competence-performance distinction and accounts for observable tendencies in the data which can neither be fully explained in terms of lexico-grammatical rules nor be reduced to mere performance phenomena. It is only by charting their always dynamic but partly conventionalized features in a representative set of naturally occurring sentences that the complex functional interplay of IOC, POC and verbs such as geben can be explained in a comprehensive way.
Kittilä, S. 2006. The anomaly of the verb ‘give’ explained by its high (formal and semantic) transitivity. Linguistics 44. 569–612.