ISIS Vol. 3

2005

Interdisciplinary Studies on Information Structure Vol. 3

UP

HUB

Approaches and Findings in Oral, Written and Gestural Language

UP | HUB

Approaches and Findings
in Oral, Written and Gestural Language

Vol. 3 (2005)

Series Editors: Shinichiro Ishihara │Michaela Schmitz
                         Universität Potsdam │SFB 632
                         Am Neuen Palais 10 │D-14469 Potsdam
                         ishihara@uni-potsdam.de;
                         mschmitz@ling.unipotsdam.de

                         Anne Schwarz
                         Humboldt-Universität zu Berlin │SFB 632
                         Sitz: Mohrenstr. 40-41
                         Unter den Linden 6 │D-10099 Berlin
                         anne.schwarz@rz.hu-berlin.de

Preface

This third volume of the working papers series *Interdisciplinary Studies on Information Structure* illustrates the diversity of approaches in play at the SFB 632 "Information Structure" (www.sfb632.uni-potsdam.de). The seven papers included offer a wide spectrum of new research findings and ongoing debates concerning focus and other information structural phenomena. Four of the papers are based on presentations at the third internal SFB workshop in Gülpe in October 2004. Contributing to this volume are members of every thematic group at the SFB, i.e. a theory developing project, typologically and diachronically oriented projects, psycholinguistic projects, and a phonetic database project.

The first paper in this volume by Elke Kasimir (Project A4: "Focus Evaluation, Anaphoricity, Discourse Coherence") discusses the reliability of the commonly used question-answer test as a focus diagnostic tool. The complications assumed by Kasimir in considering a category of *givenness* and her proposed alternative account are challenged and discussed in the following paper by Thomas Weskott (Project C1: "Contextually Licensed Non-canonical Word Order in Language Comprehension"). In the third paper, Paul Elbourne, also from the A4 project, looks at four phenomena that are particularly troublesome for theories of ellipsis and offers a new semantic analysis.

While these papers treat their subjects on the basis of English examples, the data of the fourth contribution by Ines Fiedler and Anne Schwarz (Project B1: "Focus in Gur and Kwa Languages") come from five Ghanaian languages of the Gur and Kwa language group. These languages have some morphosyntactically heavily marked focus constructions which are analyzed and diachronically interpreted. Taking a diachronic approach as well, Roland Hinterhölzl, Svetlana Petrova and Michael Solf (Project B4: "The Role of Information Structure in the Development of Word Order Regularities in Germanic") examine the interaction between information structure and word order in Old High German based on data from the Tatian translation (9th century) and find support that the finite verb form in Early Germanic distinguishes the information-structural domains of Topic and Focus.

Anke Sennema, Ruben van de Vijver, Susanne E. Carroll, and Anne Zimmer-Stahl (Project C4: "Prosody and Information Structure as Forms of Input in Second Language Acquisition") consider the effect of focus accent, word length

and position for native and nonnative perceptual processing of semantic information, as investigated experimentally. The last paper by Stefanie Jannedy and Norma Mendoza-Denton (Project D3: "Signal Parameters Connected to Prominence and Phrasing within Spoken Utterances in Different Languages") departs from classical phonetics to explore how gesture and intonation interact to structure and align information in spoken discourse, specifically through a co-occurrence of pitch accents and gestural apices.

This new volume of the series ISIS reflects the broad range of the SFB's research interests and we hope that it incites further studies in information structural phenomena in language in all its manifestations.

Shinichiro Ishihara
Michaela Schmitz
Anne Schwarz

Contents

Question-answer test and givenness: some question marks

Elke Kasimir

University of Potsdam

In order to investigate the empirical properties of focus, it is necessary to diagnose focus (or: "what is focused") in particular linguistic examples. It is often taken for granted that the application of one single diagnostic tool, the so-called *question-answer test*, which roughly says that whatever a question asks for is focused in the answer, is a fool-proof test for focus. This paper investigates one example class where such uncritical belief in the question-answer test has led to the assumption of rather complex focus projection rules: in these examples, pitch accent placement has been claimed to depend on certain parts of the focused constituents being *given* or not. It is demonstrated that such focus projection rules are unnecessarily complex and in turn require the assumption of unnecessarily complicated meaning rules, not to speak of the difficulties to give a precise semantic/pragmatic definition of the allegedly involved *givenness* property. For the sake of the argument, an alternative analysis is put forward which relies solely on alternative sets following Mats Rooth's work, and avoids any recourse to *givenness*. As it turns out, this alternative analysis is not only simpler but also makes in a critical case the better predictions.

Keywords: Focus, Givenness

1 Focus diagnostics

In order to investigate the empirical properties of focus, it is necessary to diagnose focus (or: "what is focused") in particular linguistic examples. This concerns typological study and corpus annotation, but also any attempt to understand focus from a theoretical point of view. In the following, I assume, following Jackendoff (1972), that focus corresponds to a syntactic feature, say

FOC, which influences interpretation and in various languages, among them German, Turkish, English, also the placement of pitch accents.[1,2] An example:

(1) a.   [she [only [likes [$_{DP}$ Sue]$_{FOC}$]]]

     b.   *Interpretation:* "She likes Sue and no one else."

     c.   *Intonation:* '*She only likes SUE*', with a pitch accent on Sue.

The FOC feature of the category [$_{DP}$ Sue] in (1a) triggers a particular interpretation which is depicted in (1b), and likewise constraints the intonation in the manner described by (1c). Some relevant aspects of the involved grammatical rules may be sketched as follows:[3,4]

(2)   **Focus alternatives:**

$$\llbracket [_Y \ldots X^1_{FOC} \ldots X^2_{FOC} \ldots X^n_{FOC} \ldots] \rrbracket^{FOC} =_{def} \{ \llbracket [_Y \ldots Z^1 \ldots Z^2 \ldots Z^n \ldots] \rrbracket$$
| Z^i replaces X^i_{FOC} in Y, and $\llbracket Z^i \rrbracket$ is of the same semantic type than $\llbracket X^i \rrbracket$ }

(3)   **'only' and focus:**

(i) $\llbracket only\ VP \rrbracket^w(x) = \forall m . m \in M \subseteq \llbracket VP \rrbracket^{FOC} \wedge m(w,x) \to m = \llbracket VP \rrbracket$,
                                                    for some contextually determined M, $|M| \geq 2$.

(ii) $\llbracket only\ S \rrbracket^w = \forall m . m \in M \subseteq \llbracket S \rrbracket^{FOC} \wedge m(w) \to m = \llbracket S \rrbracket$,
                                                    for some contextually determined M, $|M| \geq 2$.

---

1 Selkirk (1996) proposes that focus is expressed by a feature she calls F, but that only those features which are not dominated by other categories with an F-feature enter semantic interpretation. These "undominated" F-features she calls FOC-features. I have followed this terminology in this paper, ignoring however F-features in general.
2 An alternative view which considers focus to be a morpheme is also compatible with the discussion in this paper.
3 As usual, $\llbracket X \rrbracket$ denotes the meaning of X, X being a syntactic constituent.
4 X,Y,Z etc. are intended to range over possible natural language expressions, in some suitable sense of "possible natural language expression".

(4)   **Pitch accent placement (focus on single-word constituents):**

$[_Y$ ... $[_X$ w$]_{FOC}$ ...], where w is a phonological word, requires a pitch accent on w.

(2) implements the by now familiar notion of focus alternatives:

$\llbracket$ [likes SUE$_{FOC}$] $\rrbracket^{FOC}$ = { $\llbracket$likes Sue$\rrbracket$, $\llbracket$likes Bill$\rrbracket$, $\llbracket$likes Ann$\rrbracket$, ... }

(3) sketches a simple semantics for *only* which follows Rooth 1992 and makes use of the concept of focus alternatives as follows:

$\llbracket$She only likes Sue$\rrbracket$ = *She likes Sue* is true and for any focus alternative vp which lies in M, vp$\neq$ $\llbracket$likes Sue$\rrbracket$, *she vp's* is false.

The use of the variable M indicates that sometimes only a restricted subset of focus alternatives plays a role for the meaning of *only*, as is demonstrated here:

(5)   She invited Sue, Bill and Ann, but she only likes SUE.

This is most probably intended to express that of Sue, Bill and Ann, she only likes Sue. In contexts where no salient subset of focus alternatives is made explicit, we expect that a not-too-small set M is chosen, since such a choice makes the use of only more informative, and it is commonly assumed that more informative readings are pragmatically preferred, as long as they are relevant to the hearer. (1a), when uttered out of the blue, most probably means that Sue is the only person she likes, but still doesn't exclude that she likes her canary Tweety, unless whether she likes Tweety or not was relevant to the hearer in the utterance situation. Further below in this paper, a pragmatic constraint is proposed which imposes a precise lower bound onto the size of the set M.

(4) formulates a prosodic constraint for single-worded focused constituents which amounts for the case of (1) to the following:

in *she likes Sue*$_{FOC}$, *Sue* must carry a pitch accent.

In the general case, so the idea, there are semantic or pragmatic regularities like (2) and (3), which determine how focus influences interpretation: *focus interpretation rules*, and phonological/syntactical rules like (4) which determine how focus influences the placement of pitch accents: *focus projection rules*.[5]
If this picture is somehow on the right track, focus cannot be directly observed in languages like English, but must instead be inferred from its intonational and/or interpretative effects.[6] You hardly have any other choice than to start with some basic insights like (2)-(4), and use them to develop hypotheses for an extended range of both focus interpretation and focus projection phenomena. As an example for this methodology, consider (6), looked upon from a hypothetical initial state of knowledge represented by (2)-(4):

(6)   If she only reads BOOKS all the time, her friends will forget about her.

The sentence most likely doesn't mean (7a), as one would expect from (2)-(4), but (7b), and surely *can* mean (7b):

---

5 The term *focus projection* usually refers to the syntactic aspects of pitch accent placement only; I use it here more inclusively as subsuming also any phonological aspects of focus realization which are relevant for focus diagnostics.
6 The situation might of course be different for languages which express FOC-features by morphological and/or syntactic means.

(7) a. "If books are the only things she reads all the time, her friends will forget about her."

b. "If reading books is the only thing she does all the time, her friends will forget about her."

If (2) and (3) are roughly adequate, (7a-b) should exhibit the following FOC-marking:

(8) a. If she only reads BOOKS$_{FOC}$ all the time, her friends will forget about her.

b. If she only [reads BOOKS]$_{FOC}$ all the time, her friends will forget about her.

Since (7b)/(8b) correspond to a possible reading of (6), it is reasonable to assume that a pitch accent on the object of a VP suffices to express that the whole VP is in focus, that way giving rise to an additional hypothesis on focus projection, which will be something like the following :

(9) **Pitch accent placement (V-O-focus):**

$[_X ... [_{VP} v O]_{FOC} ...]$, where v is a verb and O its direct object, can be realized prosodically just like $[_X ... [_{VP} v O_{FOC}] ...]$.

(9) is actually an instance of the generally agreed upon projection rule *the argument projects* and its particular formulation, which involves recursion, requires for its justification a broader range of data than just (6). But it is still the case that (9), or more general formulations, are derived in the manner just described: start with observations of focus interpretation and focus projection in simple examples and use them to understand the more complicated cases.

With the same line of argument as above, using example (10), we can also derive the rule (11) which will turn out to be useful below:

(10) a.   Although she only read three BOOKS, she was well informed.

b.   *Meaning to be expected iff „books" is the focus:* „Although books were the only things she read three pieces of, she was well informed."

c.   *Preferred reading:* „Although three books were the only things she read, she was well informed."

That (10c) is a possible interpretation of (10a) suggests that the accent on books can license a FOC-marking of *three books*, which is another instance of the principle *the argument projects* and is formulated for the purposes of this paper as follows:

(11)   **Pitch accent placement (D-N-focus):**

$[_X ... [_{DP} \text{d n}]_{FOC} ...]$, where d is a determiner and n a noun, can be prosodically realized like $[_X ... [_{DP} \text{d n}_{FOC}] ...]$.

In the preceding two cases, the focus interpretation rules have been kept constant in order to derive new projection rules. But we can also go the other way around and investigate new focus-sensitive semantic/pragmatic phenomena, keeping the inventory of focus projection rules constant. One case in question is the interpretation of the sentence *that is the problem* whose interpretation seems to be sensitive to the focus marking in the statement which is the antecedent of that:[7]

---

7 The following examples are inspired by Fred Dretske's *I adviced her to steal the bicycle* (Dretske 1977).

(12) a.   SHE stole the bicycle. That is the problem.

        implication: "If someone else had stolen the bicycle, that wouldn't necessarily be a problem."

    b.   She stole the BIcycle. That is the problem.

        implication can be:

        (i) "If she had stolen something else, that wouldn't necessarily be a problem."

        or:

        (ii) "If she had done something else, that wouldn't necessarily be a problem."

According to the projection rules established so far, and (2), the following FOC-marking should be possible for (12a-b):

(13) a. SHE_{FOC} stole the bicycle

    b.   She stole [the BIcycle]$_{\text{FOC}}$

    c.   She [stole the BIcycle]$_{\text{FOC}}$

But these just give rise to the following focus alternatives according to (2):

(14) a. $[\![\text{SHE}_{\text{FOC}}\text{ stole the bicycle}]\!]^{\text{FOC}}$
$$= \{\ x \text{ stole the bicycle} \mid x \text{ an individual}\ \}$$

    b. $[\![\text{She stole [the bicycle]}_{\text{FOC}}]\!]^{\text{FOC}} = \{\ \text{she stole } x \mid x \text{ an individual}\ \}$

    c. $[\![\text{She [stole the bicycle]}_{\text{FOC}}]\!]^{\text{FOC}} = \{\ \text{she vp'ed} \mid vp \text{ a VP-meaning}\ \}$

Seemingly, the pattern is roughly as follows: *That is the problem* expresses that the preceding sentence describes the problem, whereas its focus alternatives do

not necessarily describe a problem. The following interpretation rule captures this:

(15)   **'that is the problem' and focus:**

⟦ S. That is the problem⟧ is true iff ⟦S⟧ is true, and the fact that ⟦S⟧ is true is the problem, and any state of affairs described by some m ∈ M ⊆ ⟦S⟧^{FOC}, m ≠ ⟦S⟧, isn't or wouldn't be a problem, for some contextually determined M, |M|≥2.[8]

What distinguishes *that is the problem* from *only*, and makes it especially useful for the task of focus diagnostics, is that the focus-sensitive element resides outside the sentence whose focus is to be diagnosed: whenever we want to determine the focus of a sentence which does not already contain *only*, at hand of *only*, we must first add one and that way necessarily alter the structure of the sentence under investigation. With *that is the problem*, we can leave the sentence as is.

2   Some qualifications

A not unimportant aspect of the problem of focus diagnostics lies in the fact that the set of established rules may turn out to make the wrong predictions even for those limited cases they are intended to capture. This section points to two known general limitations of the particular rules proposed so far.

The first of these concerns the determination of focus alternatives as defined in (2), has been at least in parts observed before (Rooth 1992,

---

8 It seems to depend on the context whether the focus alternatives are factual or counterfactual. In (12a), the context clearly forces a counterfactual interpretation, whereas (12b) is compatible with a reading where she actually stole something else without causing thereby any problem.

Bonomi&Casalegno 1993) and shows up whenever a generalized quantifier is in focus:[9]

(16)   Only ⟦[few STUdents]$_{FOC}$ own a bicycle].

The relevant alternative set is expected from (2) to be as follows:

(17)   ⟦[few STUdents]$_{FOC}$ own a bicycle⟧^{FOC}
            = { ⟦*dp* owns a bicycle⟧ | *dp* a possible natural-language-DP }

But then the proposition expressed by the following sentence should be part of the focus alternatives:

(18)   [Everyone who owns a bicycle, if a linguist snores, otherwise someone who doesn't own a bicycle] owns a bicycle.

But (16) just means the same as:

(19)   A linguist snores.

In other words, (19) is part of the focus alternatives relevant in (16) and is expected to show up in at least some readings of (16). It should then be possible to use (16) and thereby entail that no linguist snores.

    But the problem is more general. Since the above argument works for arbitrary propositions, not just for the proposition that a linguist snores,

    ⟦[few students]$_{FOC}$ own a bicycle⟧^{FOC}

will contain any proposition there is. But then it is just identical to

---

9 In examples like this, I attach *only* to the whole sentence, not just the DP for simplicity. Nothing depends on this, as the reader is invited to check by herself.

$$[\![[\text{few students own a bicycle}]_{\text{FOC}}]\!]^{\text{FOC}}$$

In other words, focus shouldn't make any difference in:

(20) a.   I only want that [few STUdents]$_{\text{FOC}}$ own a bicycle.

b.   I only want that [few STUdents to own a BIcycle]$_{\text{FOC}}$.

But of course, focus *does* make a difference here. The problem cannot by the way be explained away with hindsight to the contextual determination of the set M in the meaning of *only* in (3), since identical contextual influences should be operative in (20a) and (20b).

Are focus alternatives perhaps generally a bad idea, a misled intuition? I am very convinced they aren't; I am also convinced that it will sooner or later be possible to come to a reasonably precise understanding of the focus alternatives being relevant in examples like (16). My optimism stems from the fact that a rather similar problem arises in the context of the determination of the answer set of a question: Traditionally, questions like *who has a bicycle* were analyzed as either having answers like *Sue has a bicycle*, or *Sue and no one else has a bicycle* (Groenendijk&Stokhof 1997). But there is an increasing awareness that in some situations, statements like *few students own a bicycle* provide the most natural answer to *who has a bicycle*. The determination of the answer set of a question thus faces a problem which is similar to the one just described for focus alternatives: arbitrary DP's should be able to occur in answers to questions, without at the same time allowing arbitrary propositions as answers. As it happens, there seems to be an increasing interest in the conditions underlying such non-classical answer sets (Ginzburg 1996, Beck&Rullmann 1999, van Rooy 2003). And it is likely that progress in the understanding of answer sets will also advance the understanding of focus alternatives in complicated cases.

For the time being, and for the limited concern of focus diagnostics, one is however always well advised to better avoid examples where the set of focus alternatives cannot reliably be determined. And this is just what I have tried hard to do in this paper, as the reader will come to notice.

There is another, probably less severe limitation of the rules presented so far. Consider:

(21) a.   She only [bought a BOOK]$_{FOC}$

b.   She only BOUGHT a BOOK

Intended meaning in both cases: "Buying a book was all she did."

It seems that both accent patterns are possible and in more or less free variation. But of course only the second accent on book is predicted by the rules so far, i.e. (9) above. The role of the preceding accent on bought might actually be subject to debate: On one possible view, this accent is a purely phonological effect that might, for instance, depend on velocity of speech, or register. On the second possible view, this accent is input to semantic/pragmatic interpretation and indicates, say, some additional contrast (see Bolinger 1989, Féry 1993, Vallduví & Zacharski 1993, Ladd 1996:223ff. for discussion and viewpoints).

A good method to circumvent potential problems arising from such potentially mere phonological accents in focus diagnostic tasks is to always look at the minimally required accents that allow a specific reading of a sentence. As the reader will have noticed, the formulations of the rules (4), (9) and (11) have been chosen such that they determine a lower limit on the pitch accents required for a particular FOC-marking, without saying anything about the upper limit. Other authors have explicitly or implicitly followed a similar strategy.

3   Question-answer test

The preceding sections may sound rather strange to someone who is accustomed to a certain very simple method of focus diagnostics: the *question-answer test*. The question-answer test relies on an allegedly both simple and uncontroversial state of affairs, one of the "two perhaps most persistent intuitions researchers have expressed about the background–focus distinction" according to Daniel Büring (to appear), and can be stated as follows:

(22)   **Question-answer test**

If a question asks for some X (X being a syntactic category), in a direct answer to this question, the constituent which corresponds to X is focused.

What the question-answer test promises us is that we need not bother about subtle semantic or pragmatic differences, but can instead rely on simple, "objective", distributional facts: whenever a question and an intuitively "direct" answer follow each other in discourse, we know what is focused in the answer by looking at the question.

It is seemingly possible to, say, derive the projection rules (4), (9) and (11) from above in an alternative way which avoids all the tricky semantic or pragmatic issues discussed there, by just employing the question-answer test:

(23) a.   „Who does she like?" „She likes Sue!"

b.   [She [likes Sue$_{FOC}$]]

c.   *Argument:* Since „Sue" corresponds to what was asked for, we know that it is FOC-marked.

(24) a. „What does she do?" „She reads books!"

    b. [She [reads books]^{FOC}]

    c. *Argument:* Since „reads books" corresponds to what was asked for, we know that it is FOC-marked.

(25) a. „What did she read?" „She read three books!"

    b. [She read [three books]$_{FOC}$]

    c. *Argument:* Since „three books" corresponds to what was asked for, we know that it is FOC-marked.

Nice as this story sounds, the rest of this paper will show that things are not nearly such simple: examples will be discussed where the question-answer leads to the assumption of, so the claim, unnecessarily complex focus projection rules and focus interpretation rules. It will finally even turn out that these complex rules make inferior predictions in some interesting cases.

    Before all this, let's have a short look on two notions which are relevant for (22) namely that of a *direct answer to a question*, and that of *the constituent that is asked for*, in order to make the following discussion more precise. Consider to this point:

(26) a. Who does she like?

    b. She likes SUE.

    c. SHE likes SUE.

    d. She DOES like Sue! (But not so much as she likes Ann)

There is little controversy that (26b) is the direct answer we are after for the sake of the question-answer-test, whereas (26c) and (26d) should obviously not count as direct answers, but merely as indirect ones. Many people would say that (26c)

is instead a direct answer to: *Who likes whom?*, and (26d) is a direct answer to *Does she like Sue?* So one possible operational definition for the concept of direct answer is to rely on plain intuition: A sentence is a direct answer to a question when people agree that it is. This, I would guess, is the common view.

As for the question of what a question asks for, the idea is to look for the syntactic constituent in the question which "replaces" in the answer the wh-pronoun of the question. In the case of (26a-b), here repeated:

(26) a.   Who does she like?

     b.   She likes SUE!

this means that it is a DP which is asked for in (26a), and that the DP *Sue* is the suitable correspondent for it in (26b). In general:

(27)   **What a question asks for (who/what-questions):**

A question of the form *who/what VP's* asks for a DP *dp* which makes the sentence *dp VP's* true. In an answer of the form *dp VP's*, *dp* is the syntactic constituent which corresponds to what is asked for.

It should by the way not go unnoticed that the idea that what "replaces" the wh-constituent in the question is focused in the answer works well here, but not necessarily in the case of (24): There, the wh-word *what* is a DP, but what is asked for is commonly believed to be a VP. This shows to what extent the believed-to-be merely "distributional" question-answer test is actually governed by plain intuition.

We have seen now that the question-answer test promises an elegant shortcut to focus diagnostics which apparently avoids the complicated procedures of focus diagnostics sketched in the initial two sections of this paper. But it will turn out soon that the question-answer test comes with a certain prize:

it forces one two assume complicated focus projection rules which need additional pragmatic input in form of a *givenness* property: the next section demonstrates this important implication of the question-answer test. After that, the main thesis of this paper will be formulated, which amounts to the claim that this prize is too high and the question-answer test should be abandoned in its role as the authoritative device for focus diagnostics. The rest of the paper will then present arguments and facts that support this claim.

4   Givenness

Consider:

(28) Q.   Who owns a bicycle?

A.   This STUdent owns a bicycle!

What does the question-answer test say is focused in (28.A)? Under the assumption that (28.A) is a direct answer to (28.Q) in the sense of the question-answer test, and using the definition of what is asked for in (27), the answer will clearly be that *this student* is focused in (28.A). This also fits our expectations, since the accent placement in (28.A) is just as predicted by (11). All this is most likely uncontroversial. But now consider the following:

(29) Q.   As for the students: who owns a bicycle?

A.   THIS student owns a bicycle!

Following the question-answer test, one is again forced to assume that in (29.A), *this student* focused, since nothing that is relevant for the question-answer test has changed. However, whereas the accent placement in (28.A) was just as

expected, the accent on *this* in (29.A) isn't. How comes, and how to account for this?

One difference between (28) and (29) perhaps worth to be considered is that (28.Q) asks for arbitrary bicycle-owners, whereas (29.Q) is contextually understood to specifically ask for a DP which selects from students. The following revision of (27) acknowledges the fact that what a question asks for might be context-dependent:

(30)   **What a question asks for (context-dependent variant):**

A question of the form w*ho VP's* asks for a DP *dp* such that *dp VP's* is true and fits any additional contextual restrictions onto an answer to *who VP's*. In the answer *dp VP's*, *dp* is the syntactic constituent which corresponds to what is asked for.

However, using (30) instead of (27) doesn't actually change the predictions of the question-answer test for (29.A): It is still the DP *this student* which should be focused according to the question-answer test. So the question remains how it comes that (28.A) and (29.A) differ in their requirements on pitch accent placement.

There seem to be only two logical possibilities: Either fix the problem from the side of the projection rules such that these somehow predict (28.A) and (29.A) to be different, or question the results of the question-answer test and assume for (29.A) a FOC-marking which is compatible with the projection rules assumed so far. The aim of this paper is to show that the second strategy is actually viable and has some surprising advantages over the first. However, recent literature on the topic (Selkirk 1996, Rochemont 1998, Schwarzschild 1999, Büring 2003) has unanimously voted for the first strategy, so that this strategy will be looked upon first.

The basic idea which is common to the just mentioned approaches (which considerably differ in other respects) is the intuition that the crucial difference between

[this STUdent]$_{\text{FOC}}$

in (28.A) and

[THIS student]$_{\text{FOC}}$

in (29.A) lies in the fact that in the latter, but not the former, *student* is *given*, or *discourse-given*. The discourse-givenness of the noun lets the the pitch accent, so to speak, move to the left, into the only other available position, which is the determiner *this*.

A projection rule which resembles (11) (=pitch accent placement / D-N-focus), but respects *givenness*, can be formulated as follows:[10]

(31)   **Intonation for determiner-noun-focus (context-sensitive variant):**

[$_{\text{YP}}$ ... [$_{\text{XP}}$ d n]$_{\text{FOC}}$ ...], where d and n are phonological words, and d is a determiner and n a noun, *and where n is given,* is realized by pitch accent on the determiner d.

The notion of *givenness* has been around for a long time in the context of focus and information structure (see for instance Halliday 1967, Chafe 1976, Allerton 1978), but its precise context has often been left open. Somehow prominent is the idea to compare *givenness* to the notion of *familiarity* which plays a role in anaphora resolution and the licensing of indefinites, or to say that given things are *presupposed*. A third approach which has been proposed by Roger

---

10 See Selkirk 1996, Rochemont 1998, Schwarzschild 1999, Büring 2003. The rule systems proposed there are of course much more general and sophisticated.

Schwarzschild (1999) relates the *givenness* of an expression to the fact that a related linguistic expression, for instance an identical expression, has been *previously uttered in discourse*. For the case of *student* in (29.A), these three ideas would amount roughly to the following:

(32) *student* is *given* iff:

    a.   There is a set of students pre-established in the context to which *student* is anaphorically related.

    b.   *student* comes with an existential presupposition such that *this student owns a bicycle* is infelicitous instead of plainly wrong in case there are no students.

    c.   A linguistic expression which is synonymous to *student* has already been mentioned in the context.

All of these conditions seem to be fulfilled in (29.A). The question-answer test, that is (22) and (27), or alternatively (30), together with the modified projection rule (31) and one of the characterizations of *givenness* in (32a-c), therefore correctly predict the accent pattern in (29.A).

5   The main thesis

Let's summarize: after the first two sections have portrayed focus diagnostics as a rather tricky issue in languages like English, due to the fact that the syntactic FOC-marking is not directly observable in these languages, the third section offered a comfortable shortcut: just look at question-answer pairs and let the question-answer test decide on FOC-marking. The preceding section however demonstrated the cost of this move: projection rules must be assumed which are more complicated and involve reference to additional pragmatic input, namely *givenness*.

The main thesis of this paper now amounts to the claim that this prize need not be paid and is perhaps too high. It will be demonstrated that an analysis is possible which allows for simpler projection rules which do not make use of *givenness*, with equal if not superior empirical properties, if only the question-answer test is abandoned. The central idea can be illustrated at hand of the problematic example (29.A), here repeated:

(29) Q. As for the students: Who owns a bicycle?

A. THIS student owns a bicycle!

Whereas the question-answer test predicts *this student* to be in focus here, simpler projection and interpretation rules obtain if one simply assumes *this* to be in focus here, in accord to (4), the projection rule for single-word focus from the first section. This will be demonstrated in the next section.

The remaining sections make frequent use of the just mentioned competing analysis of (29.A) - focus on *this student* versus focus on *this* - in order put forward various arguments in favor of the second option: section 6 demonstrates that an alternative approach to focus in answers can be formulated which just predicts focus on *this* in (29.A) and does so without any recourse to *givenness*. Section 7 shows that focus projection rules which respect *givenness*, like (31) above, in turn require interpretation rules which also respect *givenness*. That the relevant pragmatic concept of *givenness* cannot be easily made precise will be argued for in section 8. Sections 9 and 10 finally show that projection which respect *givenness*, together with interpretation rules which respect *givenness*, even make the inferior predictions in critical examples.

6   An alternative analysis to focus in answers

The preceding section demonstrated that in order to stick to the predictions of
the question-answer test in examples like:

(29) Q.   As for the students: Who owns a bicycle?

    A.   THIS student owns a bicycle!

special amendments to the focus projection rules are necessary and have indeed
been proposed in the literature, which rely on additional pragmatic input in the
form of a *givenness* property. This section investigates the alternative option:
ignore the question-answer test and instead keep the projection rules simple. At
the core of this endeavor lies the idea to re-analyze (29.A) as involving narrow
focus on *this*:

(33)   THIS$_{FOC}$ student owns a bicycle.

The obligatory pitch accent on *this* now follows immediately from the projection
rules assumed so far, i.e. (4) (intonation for single-word focus). However, as
was explained in detail above, the FOC-marking in (33) contradicts the
predictions of the question-answer test, which consequently must be assumed to
be invalid for examples like (29.A). As a substitute, this section proposes an
alternative pragmatic rule for focus in answers to questions which just predicts
the following FOC-marking:

(28') Q: Who owns a bicycle?

    A: [This STUdent]$_{FOC}$ owns a bicycle.

(29') Q: As for the students: Who owns a bicycle?

     A: THIS$_{FOC}$ student owns a bicycle.

This alternative analysis of focus in answers will be based on focus alternatives - no separate notion of *givenness* is involved. In total, this section thus demonstrates that there is actually no real need for complicated *givenness*-related projection rules if one is willing to give up the unrestricted validity of the question-answer test as a pre-theoretic tenet.

The alternative rule for focus in answers to questions proposed now is not actually new but basically just combines the familiar approach to this very issue in Rooth 1992 with an additional constraint on focus interpretation which was apparently already present in a paper by Roger Schwarzschild (1992).[11] It goes as follows:

(34) **Focus in Answers to questions (preliminary formulation)**:

An answer A to a question Q must be FOC-marked such that $[\![A]\!]^{FOC}$ is a *minimal superset of the contextually appropriate answers to Q.*

The key idea is to explain the difference in FOC-marking between (28.A) on the one hand, (29.A) on the other hand, with the different size of the contextually appropriate answer sets of (28.Q), (29.A) resp. Whereas (28.Q) expects as an answer just *x owns a bicycle* for some arbitrary x, (29.Q) just expects an answer *x owns a bicycle*, where x is restricted to be a student. The intuition that the set of contextually appropriate answers is limited in this manner for (29.A) was already mentioned above and taken there as an opportunity to formulate (30). However, it turned out that this doesn't change the predictions of the question-

---

11 Here cited after Truckenbrod 1995. The original text was not available to the author at the time this paper was written.

answer test. Principle (34) on the other hand links the size of the contextually salient answer set to the FOC-marking in the answer by way of the alternative sets generated by the latter. It states that differences in the size of the answer sets should be reflected in the size of the focus alternatives of the resp. answers. Let's assume that these are the relevant sets:

(35) a. *contextually appropriate answers to (28.Q):*

   { x owns a bicycle | x an individual }

   b. *contextually appropriate answers to (29.Q):*

   { x is one of the students and owns a bicycle | x an individual }

   c. ⟦ *[This STUdent]$_{FOC}$ owns a bicycle* ⟧^{FOC}:

   { x owns a bicycle | x an individual }

   d. ⟦ *THIS$_{FOC}$ student owns a bicycle* ⟧^{FOC}:

   { x is a student and owns a bicycle | x an individual }

Both (35c) and (35d) are supersets of (35b), but (35d) is the smaller one. Since (34) requires the focus alternatives to be minimal, the FOC-marking which generates the smaller alternative set, which is (35d), must be chosen for (29.Q). However, since (35a) has only (35c) as a superset - (35d) being too small - the FOC-marking which generates (35a) is chosen for (28.Q). This is the idea behind (34) which will be made more precise shortly.

One point that deserves explanation and should perhaps be clarified first is how the resp. alternative sets (35c) and (35d) are actually obtained. For the case of (35c) this is rather simple: it is assumed that *this student* is a directly referring expression which just contributes an individual to the interpretation, just like *Sue* or *her*. In the case of (35d), the matter is a bit more complicated: in order to align to the analysis of *this student* as a referring expression, it is assumed here that *this* is a function from a predicate (*student* in the case of *this*

student) into an individual such that the individual referred to is in every world of evaluation the same. As a result, *this*, combined with *student*, is in fact a directly referring expression which denotes *this student* irrespective of the world of evaluation.

The following definition captures this:

(36) **Meaning and type of 'this':**

    (i)   *this N* selects for any world *w* from the individuals who satisfy *N* in *w* just one individual the speaker of *this N* intends to refer to (and, perhaps points the hearer to extra-linguistically). *this N* can be thought of as being n-way ambiguous, where n=| $[\![N]\!]$ |, and where the particular reading is contextually disambiguated at hand of the intentions of the speaker.

    (ii)   The type of *this* is the set of functions f such that i=f(w,p(w)) for every world w and 1-place-predicate intension p, i an individual, and such that p(i) holds and f is constant in its first argument.

According to this definition, the type of *this* is identical to the set of all its different readings: For the sake of focus alternatives, *this*, as used to refer to this student, has as its alternative again *this*, this time used to refer to that student, and so forth. This gives us alternatives which range over the set of all students. Notice again that this approach to the meaning and type of *this* is just consistent with the treatment of *this student* as a directly referential term in the calculation of the focus alternatives in (35c). Readers who would prefer to include the whole range of determiners: *a*, *few*, *many* etc. into the focus alternatives of *this* should consequently also include *a professor*, *few teachers*, and *many students* into the focus alternatives of *this student*. But then, the set of the focus alternatives will just explode to include any proposition, as has been shown in section 2 above. The solution presented here seems to me to be both intuitively correct and avoiding the difficulties which arise from focused generalized

quantifiers when fed into the meaning rule (2) (Def. $[\![.]\!]^{\text{FOC}}$) from the first section. [12]

Let's now return to the alternative account for focus in answers which was formulated in a preliminary fashion in (34) above. The following presents this account in a more explicit fashion and also derives it from a rather general principle of focus interpretation, called *minimize FOC-marking* here:

(37)   **Focus in answers to question (a bit more precise):**

   (i) Let Q^{ANS} be the set of contextually appropriate direct answers to Q:

   (ii) $[\![Q]\!]^{\text{ANS}} = M \subseteq [\![A]\!]^{\text{FOC}}$ for any answe r A to Q, for some suitable M.

---

12 More can be said to the meaning of *this* as described in (36) and the the resulting focus alternatives:

   (i) (36) is a simplification in that it does not distinguish between interpretation against the world of evaluation, and interpretation against the utterance context. It does furthermore not distinguish between presupposed and asserted parts of an utterance containing *this*.

   (ii) What has been described as the semantic type of *this* in (36.ii) which effectively constraints $[\![\text{this}_{\text{FOC}}]\!]^{\text{FOC}}$ is perhaps better understood as constraints resulting directly from the semantic types, and constraints on possible natural language expressions (see footnote 4 above), i.e. conservativity of determiners, which leads to the requirement that $[\![\text{this N}]\!] \in [\![\text{N}]\!]$. This still doesn't explain why the focus alternatives of *this N* are rigid designators, and perhaps they just aren't. In this case, the members of $[\![\text{Mary}]\!]^{\text{FOC}}$ probably aren't rigid designators either. A lot of technical details would change, but as far as I can see, nothing relevant to the line of argument presented in this paper.

   (iii) As an alternative to the assumption of a rather restricted semantic type of *this*, one might assume that what actually is focused in *THIS student owns a bike* is not the whole meaning of *this*, but only a part. Compare to this end *he WENT there* , where focus on the verb can optionally express focus on tense.

   All this is interesting in its own right, but should not affect any arguments presented in the text. What is crucial for the argumentation is firstly that $[\![\text{this}_{\text{FOC}} \text{ N}]\!]^{\text{FOC}} \subseteq [\![\text{[this N]}_{\text{FOC}}]\!]^{\text{FOC}}$ holds, and secondly that $[\![\text{this}_{\text{FOC}} \text{ N}]\!]^{\text{FOC}}$ restricts alternatives to members of N in a way that $[\![\text{[this N]}_{\text{FOC}}]\!]^{\text{FOC}}$ doesn't.

(38)   **Minimize FOC-marking:**

X, X being a syntactic structure inclusive FOC-marking of some sentence or text, is disambiguated to mean some *m* such that there is no grammatical structure Y in the same language which differs from X only in FOC-marking such that $Y^{FOC} \subset X^{FOC}$ and Y can be disambiguated to have the same meaning *m,* where meaning includes both propositional content and contextual restrictions. "Disambiguation" in this definition means the choice of suitable subsets M of the focus alternatives.

(38) effectively imposes a lower bound on the sets M in each of the definitions (3), (15) and (37a) such that the condition $M \subseteq [...]^{FOC}$ which is found in every of these definitions is strengthened to the condition that the resp. $[...]^{FOC}$ must be the minimal superset of M which can be expressed in that language by distributing FOC-features over the syntactic structure without changing it in other ways. Although this principle was introduced here in the context of focus in answers, it is formulated as a more general claim. And this seems to be justified. Just as an example, consider:

(39) a.   Mary [stole a BIcycle]$_{FOC}$. That is the problem.

b.   Mary STOLE$_{FOC}$ a bicycle. That is the problem.

c.   Mary stole a BIcycle$_{FOC}$. That is the problem.

From principle (38), it follows that (39a) may not get an interpretation that could also be expressed by (39b) or (39c). Consider for instance, an interpretation which is based on the following choice of M:

M={Mary stole a bicycle, Mary bought a bicycle}

Such M might be intended in a context where Mary was expected to buy a bicycle when she actually stole it. Both (39a) and (39b) are compatible with this

interpretation. But then, by principle (38), only (39b) will actually receive it, since it has the smaller alternative set. This reading is thus expected to be actually unavailable for (39a). And this just seems to fit the facts: We can (39a) use only when we want to contrast Mary's stealing the bicycle with a more general set of behavioral options. If we want just contrast Mary's stealing the bicycle with buying the bicycle, we have to use (39b).

What is so nice is that (38) can be thought of as an instantiation of the Gricean quantity maxim: If the speaker had had some $M \subseteq X^{FOC}$ in mind which could be better approximated by some FOC-marking variant Y^{FOC}, she should have chosen to utter Y. This in turn let's the hearer assume that there is no such Y.

It is hard to deny that the just presented analysis is simple and extends the basic observation of focus projection and focus interpretation in a natural and straightforward way. This section started with the proposal that (29), here repeated,

(29) Q.   As for the students: Who owns a bicycle?

   A.   THIS student owns a bicycle!

should be analyzed with a narrow focus on *this* instead of wide focus on *this student*, as was predicted by the question-answer test. It turns out now that such a re-analysis is not only possible, but perhaps even to be preferred, since no additional input in form of *givenness* must be invoked for the explanation of the accent pattern.

7   Givenness in interpretation

In the last section the possibility has been discussed that focus in answers to questions may be narrower than is usually assumed. In (29.A), here again repeated:

(29) Q: As for the students: Who owns a bicycle?

A: THIS student owns a bicycle.

the assumption of narrow focus on *this*, instead of wide focus on *this student*, contradicts the question-answer test but allows for a simpler focus projection rule. This section will argue that the simpler focus projection rule also helps to keep the interpretation rules simple. Or to put it the other way around: it will turn out that projection rules which involve *givenness* in turn require interpretation rules which also involve *givenness*. Two examples will be discussed to this end which are closely related to (29.A), the first of which being the following:

(40) Q: As for the students: Who owns a bicycle?

A: Only THIS student owns a bicycle.

implication: No other student owns a bicycle, and nothing is said about non-students.

The contribution of *only* to the interpretation of (40.A) amounts just to the claim that no other student owns a bicycle except the one denoted by *this student*. It is not excluded that there might be other, non-student owners of a bicycle. As it

turns out, exactly this meaning is predicted by (2) (Def. $\llbracket . \rrbracket^{FOC}$) and (3) (=*only* and focus) if we assume (40.A) to be the following:[13]

(41)   Only [$_S$ THIS$_{FOC}$ student owns a bicycle]

As was discussed at length in the previous section, this leads to the following focus alternatives for the embedded S:

$\llbracket$ [$_S$ [THIS$_{FOC}$ student] owns a bicycle] $\rrbracket^{FOC}$ =

{ x is a student and owns a bicycle | x an individual }

According to (3.ii) (=*only* and focus), the meaning of (40.A) is then:

(42) λ w . $\forall$ m . m $\in$ M $\subseteq$ { x is a student and owns a bicycle
            | x an individual } $\wedge$ m is true in w $\rightarrow$ m=this student owns a bicycle.

This well corresponds to what has just been described to be the meaning of (40.A): in all alternatives, the bicycle-owner is a student. Bad news for everyone who considers the complement of *only* in (40.A) to be a direct answer to the question (40.Q) in the sense of the question-answer test. Since in this case, one had to assume instead wide focus:

Only [$_S$ [THIS student]$_{FOC}$ owns a bicycle]

This of course suggests the following bigger set of focus alternatives:

(43) $\llbracket$ [$_S$ [THIS student]$_{FOC}$ owns a bicycle] $\rrbracket^{FOC}$ =

{ x owns a bicycle | x an individual }

---

13 *only* is again taken as having scope over the whole sentence in order to simplify the presentation.

According to (3.ii), the meaning of (40.A) which obtains now is:

(44) $\lambda \, w \, . \, \forall \, m \, . \, m \in M \subseteq \{ \, x \text{ owns a bicycle} \mid x \text{ an individual} \, \}$
         $\wedge \, m$ is true in $w \rightarrow m$=this student owns a bicycle.

Other than (42), (44) predicts that a reading of (40.A) is available which just says that this student is the only person who owns a bicycle. *only* doesn't range anymore exclusively over students, but over individuals in general. If one furthermore assumes that (38) (= minimize FOC-marking) above is correct, such a stronger reading is even enforced. So it seems that the assumption of narrow focus, as in (41), fits the facts much better here than the wide focus predicted by the question-answer test.

    This is not of course already a conclusive argument against the question-answer test: firstly, one could reply that the embedded *This student* owns a bicycle has indeed narrow focus on *this* in (40.A), but is not a direct answer to (40.Q) and thus need not obey the question-answer test. Since the concept of a direct answer is purely intuitive, it is hard to argue for or against such a move. Secondly, one could argue that (2) (Def. $[\![.]\!]^{FOC}$), or the deployed semantics of *only*, as given by (3.ii), is not appropriate for the cases in question: perhaps, the fact that *student* is given in the context of (40.A) influences not only the accent placement, but also either the focus alternatives, or the meaning rule of *only*. In order to save the question-answer test, one had to find a re-formulation of these interpretation rules such that they respect the *givenness* status of the constituents of the focused phrase.

    The following sketches how such a reformulation of (2) (Def. $[\![.]\!]^{FOC}$) could look like:

(45)   **Focus alternatives (variant which respects givenness):**

$\llbracket$ [$_Y$ ... X^1$_{FOC}$ ... X^2$_{FOC}$ ... X^n$_{FOC}$ ...] $\rrbracket^{FOC}$ = { $\llbracket$ [$_Y$ ... Z^1 ... Z^2 ... Z^n ...] $\rrbracket$ | Z^i replaces X^i$_{FOC}$ in Y, $\llbracket$Z^i$\rrbracket$ is of the same semantic type than $\llbracket$X^i$\rrbracket$, and for every K$_{given}$ which occurs somewhere inside of some X^i$_{FOC}$ the following holds: Any substitute Z^i for [$_X$^{i} ... K$_{given}$ ...] must have the form [$_Z$^{i} ... K$_{given}$ ...], where K$_{given}$ occupies equivalent structural positions in X^i and Z^i.}

For instance:

$\llbracket$ [$_S$ [THIS student$_{given}$]$_{FOC}$ owns a bicycle] $\rrbracket^{FOC}$

={ $\llbracket$ [$_Z$ A student] owns a bicycle $\rrbracket$ | for some suitable A and Z}

={ x is a student and owns a bicycle | x an individual }

As a result of this modification, the following two structures now generate identical focus alternatives:

[this student$_{given}$]$_{FOC}$ owns a bicycle

[this$_{FOC}$ student owns a bicycle]

(45) therefore allows one to stick to the prediction of the question-answer test that *this student* is in focus in (40.A), and at the same time obtain an interpretation which behaves as if only *this* was in focus.

   To summarize: Examples like (40) can be made consistent with the question-answer test by either claiming that *this student owns a bicycle* in (40.A) is not an answer to (40.Q) and has narrow focus on *this*, or by the assumption of more complicated focus interpretation rules which respect *givenness*.

   Let's now look at the second example to be discussed in this section:

(46) Q: As for the students: Who owns a bicycle?

A: THIS student owns a bicycle. That is the problem.

intuitive meaning: That this student owns a bicycle is the problem, and when some other student owned a bicycle, that wouldn't necessarily be a problem. Nothing is said about non-students who own a bicycle.

It again turns out that the assumption of narrow focus on *this* just gives the right meaning under the already established focus interpretation rules (2) (Def. $\llbracket . \rrbracket^{FOC}$) and (15) (=*that is the problem* and focus). The focus alternatives are of course just as before:

$$\llbracket \; [_S \; [THIS_{FOC} \; student] \; owns \; a \; bicycle] \; \rrbracket^{FOC} =$$

$$\{ \; x \; is \; a \; student \; and \; owns \; a \; bicycle \mid x \; an \; individual \; \}$$

(15) says that there is a salient subset M of these alternatives, such that any state of affairs described by some $m \in M$ obtained, this state of affairs wouldn't be a problem. Since M is a subset of the focus alternatives, under any possible choice of M, these factual or counterfactual states of affairs always involves students who own bicycles. It is thus predicted that (46.A) says nothing about bicycle-owning non-students. But this just meets the intuitive meaning of (46.A).

In an analysis which assumes wide focus on *this student* and the usual focus interpretation rules from the first section of the paper, the set of focus alternatives is considerably larger:

$$\llbracket \; [_S \; [THIS \; student]_{FOC} \; owns \; a \; bicycle] \; \rrbracket^{FOC} =$$

$$\{ \; x \; owns \; a \; bicycle \mid x \; an \; individual \; \}$$

The contextually determined subset M postulated in (15) may in this case also contain states of affairs where non-students own a bicycle, and will contain them if (37) is correct. However, (46.A) does not seem to have a reading where anything is said about non-students owning a bicycle.

As in the previously discussed example - (40) - one of the following arguments can be put forward in order to save the question-answer test: (i) (46.A) is not a direct answer to (46.Q) in the sense of the question-answer test, and in (46.A) there is a narrow focus on *this*, or (ii) the relevant focus interpretation rules are in fact sensitive to *givenness*, and the problem is just that the rules deployed here, (2) and/or (15), actually need to be revised.

However, argument (i) is much less convincing in the case of this second example, since (46.A) just feels, smells and tastes like a direct answer to (46.Q). If it wasn't, this would certainly also constitute a strong argument against the question-answer test, if not with respect to its validity, so at least with respect to its practical applicability.

So the only remaining counter-argument is (ii), which amounts to the claim that the relevant focus interpretation rules must be made sensitive to *givenness*, for instance in the way that was depicted in (45) (=focus alternatives which respect *givenness*). Since such more complicated interpretation rules can also be avoided, as has been shown above, (46.A) clearly supports the main claim of this section: that the assumption of the validity of the question-answer test leads not only to unnecessarily complex focus interpretation rules, but also to unnecessarily complex focus projection rules.

8   A closer look onto Givenness

In the previous three sections, two alternative analysis have been proposed for (29.A), here again repeated:

(29) Q: As for the students: Who owns a bicycle?

     A: THIS student owns a bicycle.

According to one of them, the accent in (29.A) falls on *this* because the otherwise preferred target *student* counts as given. The competing analysis does without any recourse to *givenness*. Considerable weight lies now on the following question: How easy is it to precisely define the relevant *givenness* property involved in the first analysis? This section presents three examples which demonstrate that finding such a definition will be far from being straightforward. These examples are closely related to the examples discussed so far. The whole problem cannot of course be captured by three examples. But it is to be expected that the problems merely increase if a larger empirical domain is being considered.

     The idea that *student* is given in (29.A) draws its initial plausibility from the fact that students have already been mentioned in the previous utterance *as for the students* in (29.Q). Without this preceding *as for the students*, the accent would fall onto *student* instead of *this* in this example. But what exactly lets *as for the students* make *student* given? Three plausible characterizations of the relevant relationship were formulated in section 4:

(32) *student* is *given* iff:

    a.   There is a set of students pre-established in the context to which *student* in *this student* is anaphorically related.

    b.   *student* comes with an existential presupposition such that *this student owns a bicycle* is infelicitous instead of plainly wrong in case there are no students.

    c.   A linguistic expression which is synonymous to *student* has already been mentioned in the context.

(32a-c) each pick out some observable relation between *students* in (29.Q) and *student* in (29.A). They furthermore enumerate the three ways in which *givenness*, or *discourse-givenness*, has been characterized in the literature: *givenness* has been related there to the properties of being *familiar*, being *presupposed*, and/or being *previously mentioned*.[14] As was already mentioned before, (32c) roughly corresponds to the formal treatment of *givenness* in Schwarzschild (1999), whose precise content can be summarized for the case of *student* as follows:

(47)   *student* is *given* iff a linguistic expression e has been uttered which translates into an open (not necessarily first-order) logical expression E with (after appropriate bijective renaming) free variables $x_1, x_2, ...,$ $x_n, (n \geq 0)$ and a free world variable w such that

$\lambda w . \exists x_1...x_n . E$ contextually entails $(\lambda w . \exists x . student(w,x)$

For the case of example (29) this amounts of course to (32c): The antecedent *student*s in (29.Q) is a suitable antecedent for *student* in (29.A), since $\lambda w . \exists X.$ students(X) logically and thus contextually entails $\lambda w . \exists x.$ students(x). Assuming that one of (32a-c) is the correct characterization of the *givenness* property which influences focus projection, the resp. version should also be able to predict de-accentuation in closely related examples. But as the examples now being discussed suggest, neither does. Consider first:

---

14 *Given material* has also been characterized in a fourth way, namely as corresponding to those parts of a sentence which are not focused, and are thus common to all focus alternatives. Analysing the *givenness* status of *student* in (29.A) in this manner of course just fits to the alternative analysis proposed above in section 6 were in (29.A) narrow focus on *this* was assumed. For the context of this paper and especially this section, the issue is however whether some *separate* property of *givenness* is needed in addition to the focus-background distinction expressed by focus alternatives.

(48) Q.   As for the students and the professors: Who owns a bicycle?

    A1.   ?THIS student owns a bicycle. And no one else!

    A2.   This STUdent owns a bicycle. And no one else!

(48.A2), with an accent on *student*, is clearly possible here and for my intuition also the only direct answer to (48.Q), whereas (48.A1) becomes appropriate only to the extent that talk is restricted to students by way of additional contextual conditions. But at the same time, all three conditions (32a-c) are satisfied, so that *student* should count as given here in any case.[15]

It is remarkable that in this example neither the alleged antecedent *students* nor the alleged anaphoric element *student* seem to have significantly changed w.r.t. example (29). The change instead consists of the presence of additional material, namely *the professors*. Anyone who insists that the difference between (29.A) and (48.A2) is due to an anaphoric relationship between *students* and *student* in (48.A2) must eventually be able to account for such indirect influences on the licensing conditions of these anaphoric links.

As it turns out, the alternative approach to focus in answers (section 6 : 37/38) which is based on the concept of focus alternatives provides an alternative explanation for the accent placement in (48.A2). In order to see this, have a look at the relevant contextually appropriate answers: in the case of (29.Q), these should be just about students who own a bike, whereas in (48.A2), answers might also be about professors who own a bike. Since according to

---

15 The supplement *and no one else* was added to the answers in (48.A2) and (48.A2) in order to signal that these are intended to provide complete utterances. This kind of explicit disambiguation was felt to be necessary since the introduction of the complex topic *the student and the professors* triggers the expectation of a more complex answer of which the sentence under discussion would then only be the first part. *And no one else* is intended to block a reading of *this student owns a bicycle* as a conversational turn which is incompletely reproduced here. I thank Thomas Wescott for hinting me to this problem.

(37), the focus alternatives of the answer have to be supersets of the answer set, (48.A2) is ruled out.

   Consider next:

(49) Q.   Since the students arrived here and brought their bicycles with them, three bicycles were stolen. Who stole the bicycles?

   A1.? THIS student stole the bicycles.

   A2. This STUdent stole the bicycles.

For my very impression, (49.A1) is only appropriate if the speaker of the question *who stole the bicycles* is in the context understood to believe that one of the students must have stolen the bicycle. In a more neutral situation, only (49.A2) is appropriate.

   On the other hand, *student* counts again as *given* according to (32a-c) and is therefore predicted to be de-accented.[16] And again, the alternative approach to focus in answers (section 6: 37/38) just predicts the correct accent placements here: If the contextually salient answer set excludes non-student thieves, (49.A1), the answer with the smaller set of focus alternatives is predicted. If the contextually salient answer set is more general, the FOC-marking found in (49.A2) is predicted.

   Consider finally:

(50) Q.   As for the students: Who owns a bicycle?

   A1.?? TWO students own a bicycle.

   A2.   TWO STUdents own a bicycle.

---

16 Condition (32a) is only satisfied if *this student* is one of the students referred to by *the students,* but this is certainly a salient reading of (49).

It is again the answer with an accent on supposedly given material which is to be preferred in this case, i.e. (50.A2). Interestingly, (50.A1), the answer where *students* is de-accented, is intuitively understood as an answer to:

(51)   How many students own a bicycle?

Assume for the moment that the FOC-marking in (50.A1) was as follows:

(52) TWO_{FOC} students own a bicycle.

That (50.A1)/(52) is well understood as an answer to (51), but not easily understood as an answer to (50.Q) could be explained by the assumption that *two* actually denotes a numeral which can only be replaced by other numerals, not by *this*, *few*, *many*, *every* or the like. In this case the relevant focus alternatives were as follows:

(53) $[\![\text{TWO}_{\text{FOC}} \text{ students own a bicycle}]\!]^{\text{FOC}}$

$$= \{ \text{ n students own a bicycle} \mid \text{n a cardinal} \}$$

But this set is only a superset of the answer set of (51), not one of (50.A1), even if the latter was restricted to ask for students. The alternative theory of focus (section 6: 37/38) can therefore at least explain why (50.A1) is infelicitous. It can of course not readily explain the peculiar accent pattern in (50.A1). Perhaps, a better understanding of focused generalized quantifiers and the alternative sets generated by them will provide for a satisfying analysis of these examples (see the discussion in section 2). It is on the other side not clear how a *givenness*-based explanation for this example should look like.

Three examples have been presented now which present each an empirical problem for the notion of *givenness* as it is usually understood. It has been

shown that these examples are at the same time amenable to an analysis in terms of focus alternatives, at least to a certain extent. The range of de-accenting phenomena which have been explained with *givenness* in the literature is of course much bigger, so that the observations just presented merely scratch the surface of the problem. They nevertheless hopefully suffice to show that the task to pin down the alleged *givenness* property such that it allows for reliable predictions will not be a straightforward task. This imposes a considerable burden on the proponents of a *givenness*-based account and demonstrates that the complications induced by projection rules and interpretation rules which respect *givenness* are far from being trivial.

I would like to add that when working on *givenness* in the second half of the year 2004 and discussing various issues with colleagues, I had the very impression that the validity of the concept of *givenness* was never put to scrutiny simply because everyone takes it for granted that one cannot do without it anyway. If this was true, counterexamples were of course not of so much interest: they just indicated work that still has to be done. I however hope to present convincing arguments in this paper that the situation actually isn't this way: especially if one is willing to abandon the unrestricted validity of the question-answer test for focus diagnostics, alternative elegant ways show up for the analysis of focus and accent placement which do not make use of *givenness*. Whether a precise and empirically satisfying characterization of *givenness* is possible therefore may turn well out to be a crucial factor in deciding between theories of focus.

9   Focus projection and constituency

Two competing analysis have been under discussion for *this student* in (29.A):

(29) Q.   As for the students: Who owns a bicycle?

   A.   THIS student owns a bicycle!

which can be depicted as follows:

   (i)   [this student$_{given}$]$_{FOC}$
   (ii)   [this$_{FOC}$ student]

These two analysis have been compared from different perspectives, and the second analysis has been claimed to be superior because of its greater simplicity given that it doesn't involve *givenness* and related complications. However, no examples have been discussed so far where the empirical predictions of both approaches clearly and unanimously diverge. The last two sections of this paper are dedicated to such examples.

   The structures (i) and (ii) are hard to distinguish because they trigger the same accent pattern and also the same interpretation. The latter of course only with suitable stipulations: the modified rule (45) for the computation of focus alternatives in chapter 7 basically states that given material, although being focused, just behaves like unfocused material for the sake of interpretation.

   It is however possible to construct examples where either accent placement or interpretation are expected to differ according to the two competing lines of analysis, even if rules like the just mentioned (45) are deployed. One structure with this desirable property is:

(54)   [$_{VP}$ v [$_{DP}$ d N$_{given}$]]$_{FOC}$

In order to re-analyze this structure such that given material is actually not part of the focus, *two* FOC-features must be assumed:

(55)   [VP vFOC [DP dFOC N]]

As it turns out, for this latter structure the projection rules predict two accents, one onto the verb and the other on the determiner, by projection rule (4) (=focus on single words). For (54) on the other hand, one accent on the determiner d should suffice according to the projection rules (9) (V-O focus) and (31) (D-N focus).

A test situation where one would expect a structure like (54) consists of an answer to a question which asks for a VP, plus additional contextual material that makes N given. In order to construct such a test case we first need to know what kind of questions ask for a VP - up to now, we have always used questions which ask for a DP. The following seems to be widely accepted in the literature:

(56)   **What a question asks for - VP meanings:**

The question *what did x do?*, $[\![x]\!]$ an individual, asks for a VP-meaning.

As an example, take:

(57) Q: What did Mary do?

A: She [stole a BIcycle]$_{FOC}$! That's the problem.

implication: If she did something else, that wouldn't necessarily be a problem. Nothing is said about someone else doing something.

The pitch accent in (57.A) is well explained by (56) in combination with (9), the focus projection rule for VP's, and (11), the focus projection rule for DP's with no given constituents. That the VP is in focus is independently confirmed by the semantic effect of that's the problem: The focus alternatives of (57.A) are:

{ she *vp* | *vp* a VP-meaning }

According to (15), the meaning rule of *that is the problem*, (57.A) should express that some significant subset of these focus alternatives describe something that isn't a problem. This just seems to fit the intuitive meaning.

Now consider the following context:

(58) Q: As for the bicycles: what did Mary do?

A: She stole this bicycle.

In this context, the question asks for a VP-meaning, and *bicycle* is at the same time *given*. The answer thus just instantiates (54):

(59) She [$_{VP}$ stole [this bicycle$_{given}$]]$_{FOC}$

Now it seems as if (58.A), to the extent that one is willing to de-accent *bicycle* in the first place, strongly prefers the following accent pattern:

(60) She STOLE THIS bicycle.

The accent on the verb is certainly indispensable here. But this just fits a re-analysis along the lines of (55):

[$_{VP}$ stole$_{FOC}$ [this$_{FOC}$ bicycle]]

The analysis depicted in (59) which follows (54) would instead predict that the accent on the verb can be omitted. So it seems that in this case, the re-analysis along the lines of (55) is not only simpler, but also makes the better prediction. The following two examples do not use the question-answer test, but instead *only* and *that is the problem*:

(61)   As for the bicycles: Mary only stole THIS bicycle.

In this case the competing analyses are the following:

(62) a.   Mary only [stole THIS bicycle~given~]~FOC~.

    b.   Mary only stole THIS~FOC~ bicycle.

I again assume for simplicity that *only* has actually scope over the whole sentence for the sake of interpretation. In order to be fair, the focus alternatives in (62a) are computed with the help of (45) from section 7 above such that they respect *givenness*. In that case (roughly):

(63)   ⟦ [Mary [stole this bicycle~given~]~FOC~] ⟧^FOC^

            ≈ ⟦ [Mary stole~FOC~ this~FOC~ bicycle] ⟧^FOC^

            ≈ { Mary (does/did/will do) x with y | x an activity, y a bicycle }

For (62b) the set of focus alternatives is considerably smaller:

(64)   ⟦Mary stole this~FOC~ bicycle⟧^FOC^ = { Mary stole x | x a bicycle }

The crucial difference becomes clear: The alternatives in (62a) can involve different verb meanings, where the verb in (62b) is fixed in all alternatives to be *stole*.

        This difference should influence the contribution of *only*: According to (62a), (61) is predicted to mean that stealing this bicycle is the only activity that Mary performs which is directed towards one of the bicycles. According to (62b), (61) is predicted to mean that this bicycle is the only one that Mary stole. Nothing is said about any non-stealing activities from the side of Mary. It seems

to be obvious to me that the second interpretation is the only one which is available for (61).

The same line of argument can be performed at hand of *that is the problem*:

(65)   As for the bicycles: Mary stole THIS bicycle. That is the problem.

with the competing analyses:

(66) a.   Mary [stole THIS bicycle$_{given}$]$_{FOC}$.

    b.   Mary stole THIS$_{FOC}$ bicycle.

The focus alternatives are the same as above. According to (66b), (67) means that if Mary stole or had stolen some other bicycle, that wouldn't necessarily be a problem. Nothing is said about other activities from the side of Mary. According to (66b) however, (68) means that if Mary had done something with one or the other bicycles, that wouldn't necessarily be a problem. It seems again to be obvious to me that the first interpretation is the only one which is available for (66).

10 Focus and PP-Extraction

This concluding section further extends the discussion of the preceding section. A famous example by Lisa Selkirk will be discussed which apparently speaks against the results from the last section. It however turns out that this example can and must be re-analyzed. Under this re-analysis, it fits very well to the examples obtained so far, and the main thesis defended in this paper.

Consider first the following VP:

(69)   [$_{VP}$ stole THIS bicycle]

As was discussed at length above, according to a *givenness*-based analysis, the pitch accent on *this* should be able here to license focus on the VP in suitable circumstances, since it instantiates the following structure:

(54)   [$_{VP}$ v [d N$_{given}$]]$_{FOC}$

According to the alternative line of analysis proposed in this paper, (69) just licenses narrow focus on the determiner *this*:

(55)   [$_{VP}$ v [d$_{FOC}$ N]]

The last section presented some empirical observations which indicated that the structure depicted in (54) is actually not available.

Basically the same situation occurs with nominal phrases which take a PP-complement. In a configuration [d N PP]$_{FOC}$, the accent is commonly assumed to fall by default onto the complement of the preposition, as is demonstrated here:[17]

(70) Q: What did Mary buy?

    A: Mary bought [a book about BATS]$_{FOC}$

Consider now the VP in:

(71)   Mary [$_{VP}$ bought a BOOK about bats]

---

17 This is again an instance of the projection principle *the argument projects* - see section 1.

According to a *givenness*-based analysis, the pitch accent on *book* should again be able to license focus on the whole VP in suitable circumstances, since it instantiates the following structure, which closely resembles (54) cited above:

(72) $[_{VP}$ v $[_{DP}$ d N $PP_{given}]]_{FOC}$

According to the alternative line of analysis proposed in this paper, (69) at most licenses narrow focus on *a book*:[18]

(73) $[_{VP}$ v $[\![$d N$]\!]_{FOC}$ PP$]]$

As it turns out, (71) is actually a famous example by Lisa Selkirk (1996). Selkirk adopts for (71) the analysis depicted in (72) and consequently claims that the whole VP can be focused in this example in any context where *about bats* is *given*. Somewhat surprisingly in light of the results from the previous section, there is evidence that the verb *bought* can indeed well be part of the focus licensed by a pitch accent on *book*:

(74)   Mary only bought a BOOK about bats.

This sentence can well mean that the only thing that Mary did with respect to bats was to buy a book about them. This indicates that the relevant alternative set allows for variation in the resp. activity that has to do with bats:

(75) $[\![$Mary [bought a BOOK [about bats]$_{given}]_{FOC}]\!]^{FOC}$

        $\approx$ { Mary (did/does/will do) x with y about bats
                | x an activity, y a suitable NP-meaning }

---

18 I assume here for concreteness that [d N] form a DP-internal constituent. If not, the following arguments have to be adjusted as to exclude the determiner from narrow focus.

At the same time, an analysis according to (73) is out, since in that case, the relevant set of focus alternatives was too small:

(76)   ⟦Mary bought [a BOOK]$_{FOC}$ about bats⟧^{FOC}

        ≈ { Mary bought x about bats | x a suitable NP-meaning }

This seems to be a strong empirical argument for projection rules which respect *givenness*. However, as it happens, (71) and (76) perhaps do not actually instantiate the syntactic structure tacitly assumed so far, for consider:[19]

(77)   Mary only bought a BOOK yesterday about bats.

In (77), the PP *about bats* has most likely been extraposed. It thus could well be that in (71) and (76) the PP has been extraposed too. The following illustrates this option:

(78)   Mary [$_{VP}$ bought [a book t$_i$]]$_{FOC}$ [about bats]$_i$

This is the analysis I want to propose for this example. I agree with Selkirk's analysis that the VP is focused in this examples. That *about bats* is ignored for the sake of focus projection I however attribute this to that fact that it has been extraposed. The relevant rule could be stated as follows:

(79)   Extraposed material, and the trace it leaves, is ignored in focus
        projection.

What kind of example could help now to decide between an analysis based on extraposition and (79) on the one hand, and an analysis based on *givenness* and

---

[19] Since several reviewers doubted this: the following example has been judged fine in appropriate contexts, and simuilar examples are discussed in Guéron (1980).

focus projection rules which respect *givenness* on the other hand? As it turns, out, not every [d N PP] structure allows extraposition, for consider:

(80) a. * Bill drank a GLASS yesterday of beer.

   b. * Bill met the KING yesterday of Belgium.

Examples based on such DP's should not be amenable to an analysis according to extraposition. Consider to this end:

(81)   Bill drank a GLASS of beer. That is the problem.

   implication: If Bill drank some other amount of beer, that wouldn't be a problem. Nothing is said about activities other than drinking beer.

(82)   Bill met the PREsident of Belgium. That is the problem.

   implication: If Bill met some other representative of Belgian, that wouldn't necessarily be a problem. Nothing is said about activities other than meeting representatives of Belgian.

The analysis based on *givenness* would still predict that the VP can be in focus in these examples. The focus alternatives should therefore include other activities than drinking or meeting someone, as is sketched here:

   { Bill (did/does/will do) x with y of beer
                           | x an activity, y a suitable *np*-meaning }

   { Bill (did/does/will do) x with y of Belgian
                           | x an activity, y a suitable *np*-meaning }

The alternative analysis predicts a narrow focus on the NP in these cases, where the activity is fixed to be drinking, meeting someone, resp.:

{ She drank x of beer | x a suitable *np*-meaning }

{ She meet x of Belgian | x a suitable *np*-meaning }

The intuitive meanings of (81) and (82) clearly favors the second analysis.

The very issue of the interaction between extraposition and information structure clearly deserves a much more thorough investigation than could be done here. In an still ongoing investigation undertaken by the author and Kepa Joseba Rodriguez into the semantics and pragmatics of extraposition, one which currently concentrates on Basque and Turkish, there is ample evidence that extraposition interacts in rather complex ways with information structure. This supports earlier observations on extraposition in English for instance by Jacqueline Guéron (1980). There is by the way some initial evidence that at least in Turkish and Basque, extraposed material can be "backgrounded" in some sense which is well distinguishable from just not being in focus. So it may turn out after all that the semantics and pragmatics of extraposition is the very location where an independent pragmatic concept of *givenness* or so is operative. The intuition that in examples like (71) from Lisa Selkirk some orthogonal pragmatic input is involved in pitch accent placement might therefore turn out to be still true in the end.

11 A final remark on focus alternatives

Throughout this paper, focus alternatives have been used as a means to describe the semantic/pragmatic impact of focus. It has nowhere been claimed that, say, focus alternatives somehow have to lie at the heart of the correct semantic approach to focus, and an implicit claim to this end is also certainly not intended from my side. But the descriptive use of focus alternatives is though not totally innocent: if it turned out that some considerably simpler and less expressive notion of *givenness* would suffice to derive the semantic/pragmatic

effects of focus, as the meaning of *only* or *that is the problem*, this might threaten the line of argument presented in this paper: opponents could argue that *givenness* can replace focus alternatives and than claim that latter be furthermore even worse than *givenness*. A discussion of this interesting issue was unfortunately far beyond the scope of this paper; I can only say that I am personally rather convinced that a concept of *givenness* like the one just sketched can never be formulated; articles which are related to this topic include Schwarzschild 1997, Geurts&van der Sandt 2004 and Jäger 2004.

12 Acknowledgments

References

Allerton, D.J. (1978). The Notion of "Givenness" and its Relations to Presupposition and to Theme. Lingua 44: 133-168.

Beck, S. & Rullmann, H. (1999). A flexible approach to exhaustivity in questions. Natural Language Semantics 7: 249-298.

Bonomi, A. & Casalegno, P. (1993). Only: Association with focus in event semantics. Natural Language Semantics 2: 1-45.

Büring, D. (1997). The 59th Street Bridge Accent. On the Meaning of Topic and Focus. London: Routledge.

Büring, D. (2003). Focus Projection and Default Prominence. In: Valeria Molnar and Susanne Winkler (eds.): Proceedings from the Symposion `Informationsstruktur -Kontrastiv'. Mouton De Gruyter.

Büring, D. (2004). Focus Suppositions. Theoretical Linguistics 30(1): 65-76.

Büring, D. (to appear). Intonation, Semantics and Information Structure. In: Gillian Ramchand and Charles Reiss (eds.): Interfaces.

Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view: 25-55. In: C. Li (ed.): Subject and Topic. New York: Academic Press.

Chapman, S. (1998). Accent in Context. The Ontological Status and Communicative Effects of Utterance Accent in English. Peter Lang Publishing.

Dretske, F. (1972). Contrastive Statements. Philosophical Review 81: 411-437.

Dretske, F. (1977). Referring to Events. Midwest Studies in Philosophy 2: 369-378.

Féry, C. (1993). The Tonal Structure of Standard German. Tübingen: Max-Niemeyer.

Geurts, B. & Sandt, R.v.d. (2004). Interpreting focus. Theoretical Linguistics 30: 1-44.

Ginzburg, J. (1996). Interrogatives: Questions, Facts and DialogueThe Handbook of Contemporary Semantic Theory. Oxford: Blackwell.

Groenendijk, J. & Stokhof, M. (1997). Questions. In: von Benthen, J. and ter Meulen, A. (eds.): Handbook of Logic and Language.

Gussenhoven, C. (1983). Focus, mode and nucleus. Journal of Linguistics. 19: 377-417.

Gussenhoven, C. (1992). Sentence accents and argument structure. In: I.M. Roca (eds.): Thematic Structure: Its Role in Grammar. New York: Foris.

Guéron, J. (1980). On the Syntax and Semantics of PP Extraposition. Linguistic Inquiry 11: 637-678.

Halliday, M.A.K. (1967). Notes on transitivity and theme in English. Journal of Linguistics 3: 199-244.

Jackendoff, R.S. (1972). Semantic Interpretation in Generative Grammar. Cambridge: MIT Press.

Jäger, G. (2004). Alternatives or presuppositions? A comparison of the Background-Presupposition Rule with Alternative Semantics. Theoretical Linguistics 30(1): 111-122.

Kaplan, D. (1978). DThat. In: Peter A. French, Theodore E. Uehling, jr. and Howard K. Wettstein (eds.): Contemporary Perspectives in the Philosophy of Language. Minneapolis: University of Minnesota Press.

Kaplan, D. (1978). On the Logic of Demonstratives. Journal of Philosophical Logic 8: 81-98.

Kratzer, A. (2004). Interpreting focus: Presupposed or expressive Meanings? A comment on Geurts and van der Sandt. Theoretical Linguistics 30(1): 123-136.

Krifka, M. (1991). A Compositional Semantics for Multiple Focus Constructions. In: Joachim Jacobs (ed.) Informationsstruktur and Grammatik: Joachim Jakobs .

Krifka, M. (1992). For a Structured Meaning Account of Questions and Answers. In: Frey, Caroline and Sternefeld, Wolfgang (eds.): Audiatur vox sapientiae. Berlin: Akademie Verlag.

Krifka, M. (1995). The Semantics and Pragmatics of Polarity Items. Linguistic Analysis 25: 209-257.

Kuno, S. (1987). Functional Syntax: University of Chicago Press.

Prince, E.F. (1981). Toward a taxonomy of given/new information: 223-255. In: P. Cole (ed.) Radical Pragmatics. NPYP: Academic Press.

Rochemont, M. (1997). Phonological Focus and Structural Focus 29. In: P. Culicover and L. McNally (eds.): The Limits of Syntax, Syntax and Semantics.

Rooth, M. (1985). Association with Focus: Amherst.

Rooth, M. (1992). A Theory of Focus Interpretation. Natural Language Semantics 1(1): 77-116.

Rooth, M. (1996). Focus: 271-297. In: Lappin Shalom (ed.): The Handbook of Contemporary Semantic Theory: Blackwell.

Rooth, M. (1996). On Interface Principles for Intonational Focus. In: T. Gallowayand S. Justin (ed.): Proceedings From Semantics and Linguistics Theory VI. Ithaka: Cornell University.

Rooy, R.v. (2003). Questioning to Resolve Decision Problems. Linguistic and Philosophy 26: 727-763.

Rooy, R.v. & Schulz, K. (2003). Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation. Netherlands: Kluwer.

Schwarzschild, R. (1992). The constructiveness of associated foci.

Schwarzschild, R. (1997) Why some Foci must Associate. Rutgers University.

Schwarzschild, R. (1999). GIVENness, AvoidF and Other Constraints on the Placement of Accent. Natural Language Semantics 7(2): 141-177.

Selkirk, E.O. (1996). Sentence Prosody: Intonation, Stress and Phrasing. In: John A. Goldsmith (ed.): The Handbook of Phonological Theory. London: Basil Blackwell.

Selkirk, E. (1984). Phonology and Syntax: The Relation between Sound and Structure. Cambridge, MA: MIT Press.

Truckenbrodt, H. (1995). Phonological Phrases: Their Relation to Syntax, Focus and Prominence. Doctoral dissertation.

Elke Kasimir
Universität Potsdam
SFB 632 „Informationsstruktur"
Postfach 601553
14415 Potsdam
Germany
elke.kasimir @catmint.de

Stop Bashing Givenness!

A Note on Elke Kasimir's "Questions-Answers Test and Givenness"[*]

Thomas Weskott

Universität Potsdam

Elke Kasimir's paper (in this volume) argues against employing the notion of Givenness in the explanation of accent assignment. I will claim that the arguments against Givenness put forward by Kasimir are inconclusive because they beg the question of the role of Givenness. It is concluded that, more generally, arguments against Givenness as a diagnostic for information structural partitions should not be accepted offhand, since the notion of Givenness of discourse referents is (a) theoretically simple, (b) readily observable and quantifiable, and (c) bears cognitive significance.

Keywords: Givenness

In her paper "Questions-Answers Test and Givenness: Some Question Marks" (this volume), Elke Kasimir argues against the question-answer test for focussed constituents on the ground that it involves the assumption of "(...) unnecessarily complex focus projection rules and focus interpretation rules" (Kasimir, 2005, p. 13)[1]). In addition to their allegedly unnecessary comlexity, these rules are incriminated for the fact that they "(...) need additional pragmatic input in form of a *givenness* property (...)" (pp.15, 18f.; Kasimir's emphasis), the addition of which the author claims to be too high a price to pay (*ibd.*). Furthermore, Kasimir claims that an alternative analysis which abandons Givenness as an ingredient in the explanation of accent placement makes better predictions for a

---

[*] Thanks to Elke Kasimir for letting me comment on her paper, as well as for clarifying discussion. The usual disclaimer holds.

[1] Henceforth, all quotes refer to that paper unless indicated otherwise.

certain class of examples.

While I do not intend to comment on the last claim, and lacking a commonly agreed upon metric for the complexity of focus projection and/or interpretation rules that would put me in a position to grapple with the first claim, I will only deal with the second, the argument against the usefulness of Givenness in determining accent assignment. But before doing so, a disclaimer might be in place. I assume that what I have to say is neutral with respect to Kasimir's first claim, i.e. I assume that my stance on the theoretical and empirical usefulness of the notion of Givenness is independent of that of the question-answer test. Furthermore, I take it that none of the following takes sides as regards the issue of which theory of focus interpretation is to be preferred (e.g., localist vs. globalist accounts, cf. Jäger (2004)). And finally, I will not try to propose a fleshed-out Givenness-based account of accent assignment or focus interpretation; for the latter, the reader is referred to Sauerland (2004), for the former, to Wagner (2005).

Kasimir argues that in order to make use of the question-answer test as a diagnostic for accent placement, one has to assume focus projection rules and focus interpretation rules which are said to be unnecessarily complex . As a further and apparently even more severe drawback, she claims that these rules need additional pragmatic input from a Givenness property, and that to assume such additional input is too high a price to pay.

My argument against this position is going to run down the following line: Firstly, I will show that Kasimir's critical example holds Givenness in store at a relatively low price—in fact, that Givenness is for free in this case. Secondly, I will argue that it is unclear exactly how Kasimir's alternative strategy, i.e. to account for the accent placement in this example by recourse to alternative sets plus a pragmatic rule that restricts the answer set, should work *without* appealing to Givenness, and that it indeed necessitates appeal to Givenness just

as much as the analysis it purports to supersede. I conclude that the argument for the alternative account *and* the argument against the Givenness-based account are equally inconclusive.

My second point will be more general and is concerned with the theoretical status of the notion of Givenness of discourse referents in determining information structural partitions. I will point out that, its shortcomings already revealed by Reinhart (1981) notwithstanding, it is indispensable as a heuristic for detecting focus-background and topic-comment structures, and that its theoretical as well as its empirical value lie in its fruitfulness for a cognitive understanding of the role information structure plays in natural language.

The pair of examples playing the central role in Kasimir's argument against a Givenness-based account of accent placement are the following (I stick to Kasimir's numbering for convenience): [2]

(28) Q. Who owns a bicycle?

    A. [This STUdent]$_{FOC}$ owns a bicycle.

As regards (28), Kasimir notes that the question-answer test (in its context-sensitive variant) correctly predicts the focus of the answer to be *this student*, which ultimately yields the correct accent assignment . This example is contrasted with the following:

(29) Q. As for the students: who owns a bicycle?

    A. THIS student owns a bicycle.

---

[2] For what follows, I will assume for the sake of argument that Kasimir's definitions for accent placement and the question-answer test as given in the first three sections of her paper are correct—by which statement I do not mean to imply that they are not, but rather that I will take them for granted.

The issue Kasimir points to is "(...) that (28.Q) asks for arbitrary bicycle-owners, whereas (29.Q) is contextually understood to specifically ask for a DP which selects from students." (p.16). The question is, then, how this difference can be linked to the differences in requirements on pitch accent placement. Kasimir offers two logical possibilities: one can either try to account for the differences between (28.A) and (29.A) by adjusting the rules for focus projection, or, and that is Kasimir's choice, by sticking to the projection rules assumed so far and fix the problem instead by assuming a different interpretive mechanism (alternative sets plus a restriction on the set of contextually salient answers). The relevant sets and the respective focus-background structures of the answers look like this:

(35) a. *contextually appropriate answers to (28.Q):*

   { x owns a bicycle | x an individual }

   b. *contextually appropriate answers to (29.Q):*

   { x is one of the students and owns a bicycle | x an individual }

   c. $[\![$[This STUdent$]_{FOC}$ owns a bicycle $]\!]^{FOC}$

   { x owns a bicycle | x an individual }

   d. $[\![$ THIS$_{FOC}$ student owns a bicycle $]\!]^{FOC}$

   { x is a student and owns a bicycle | x an individual }

Let us first consider the first alternative, which Kasimir rejects. The reason for this rejection is that it involves the adjustment of focus projection rules so as to respect the property of Givenness (cf. her rule (31), p.17), which is said to "(...) involve reference to additional pragmatic input (...)" (p.18).

Let me comment on that last point first. No matter whether one assumes a dynamic or a classical interpretation procedure, I take it that establishing the

relation between the wh-pronoun restricted by the *as for*-phrase and the subject of the answer involves no *additional* input whatsoever, be it semantic or pragmatic, since the anaphoric relation between the two sets of discourse referents is *just there*. So Givenness has not to be stipulated or derived for any price, it can be had for free. This point can be made even stronger if the question is paraphrased by the following variant featuring a d-linked wh-pronoun, which does the same (at least for the issue at hand) as the *as for*-phrase, viz. to restrict the domain of the wh-element to the extension of the noun:

(29') Q. Which student owns a bicycle?

     A. THIS student owns a bicycle.

The next question is how the focus projection rules can be made sensitive to Givenness or the alternative property proposed by Kasimir. According to her, this alternative works as follows: firstly, the set of contextually appropriate answers are computed, and so are the alternative sets generated by the respective answers; the minimal set matching both restrictions then decides which is the correct focus-background structure, which in turn yields the correct accent assignment for the answer. On the Givenness account, the projection rules have to be made sensitive to the fact that the restrictor of the complex demonstrative *this student* is mentioned, presupposed, or inferable from context (cf. p.18).

My worry here is the following: it is crucial to Kasimir's argument that the way in which the complex demonstrative *this N* contributes to the computation of the alternative sets containing it does not depend on any kind of Givenness (on pain of begging the question). So obviously it is not sufficient to determine the meaning and the semantic type of a demonstrative of the form *this N* (as Kasimir does on p.24), but one also has to pin down the information structural properties of such phrases, more specifically: whether a phrase of the form *this N* carries the presupposition that there exists an N. Whatever the theoretical

stance on this question (s. Lepore & Ludwig (2000) for a discussion), I think that either answer will fit uncomfortably with Kasimir's alternative proposal. That is because if we assume that a complex demonstrative of the form *this N* has roughly the logical form

[the x : x is this and x is N]

and further assume that the quantifier *the x* carries an existential and a uniqueness presupposition which enters into the computation of the alternative set of an utterance containing it, Kasimir's alternative proposal will become corrupt because it begs the question of the role of Givenness. On the other hand, if we assume that no such presupposition is present in the computation of the alternative sets, the question arises where these sets come from in the first place. If I understand her correctly, Kasimir argues that in order to arrive at the right kind of alternative sets, she makes use of a notion of Givenness she considers uncontroversial, namely one that characterizes as given something which is "(...) corresponding to those parts of a sentence which are not focused, and are thus common to all focus alternatives." (p. 34, fn.14).

Now it seems obvious to me that this way of reasoning simply reverses the order of explanation (a focus-background structure is taken for granted in the construction of the alternative sets, instead of derived at by taking the latter and the context into consideration), and hence puts the argument for Kasimir's alternative account in jeopardy of becoming circular.

Either way, I fail to see in which way Kasimir's alternative account can make sense of the minimization of the superset of the set of contextually appropriate answers and the alternative sets *without* relying on some notion of Givenness, be it one that comes from the presupposition of the question, or one stemming from the presupposition of the complex demonstrative, or both. Her way of deriving at the alternative sets by taking a focus-background structure for granted

is at best inconclusive, at worst circular. Moreover, given the straightforward anaphoric relation between the set denoted by the question and the set denoted by the focussed part of the answer, I do not see a reason for her reluctance to employ a notion of Givenness in the derivation of accent placement, its price not being, as I have argued, high at all.

The reason for discussing Kasimir's claims with respect to the notion of Givenness—apart from the formal details criticised above—is as more general one: critical evaluations of that notion which are similar in spirit to that of Kasimir's (e.g. in Reinhart, 1981, and, more recently, Krifka, 2004, a.o.) have shown it to be neither sufficient, nor even necessary for the determination of focus-background and topic-comment structures and, accordingly, accent assignment, in various cases. While I certainly do not intend to question these findings, I want to raise the issue whether they countervail against the use of that notion in information structure research in general. My answer to that question is in the negative, and it is for the following three reasons.

Firstly, I consider the notion of Givenness, for all its shortcomings in some special cases, to be the simplest observable information structural property of discourse referents. Although the notion of anaphoricity is not unproblematic either, having an explicit antecedent in context (i.e., being mentioned) is a feature of natural language expressions that can be detected even by non-linguists. That alone would, of course, not make it theoretically interesting.[3] But, secondly, properties such as the systematicity of distance of antecedent and anaphoric element on the one hand, and form of expression (s. e.g. Ariel, 1990) on the other, as well as the systematic recurrence of given elements in referential (or topic) chains illustrate the explanatory potential of the notion of Givenness in

---

[3] Although it certainly is a virtue in e.g. instructing naive annotators how to detect information structural properties of parts of speech in corpora, which would be even more troublesome an endeavour if one were solely to rely on notions like Relevance, Contrast, Aboutness, etc.

an exemplary fashion, since both of these properties are quantifiable and hence can be put to empirical test. This brings me to the third point: not only can the systematicities of Givenness explain certain properties of texts, they moreover have been shown to be highly effective in the actual production and perception of natural language (s. Garrod & Sanford, 1994, for an overview). To give but one recent example: the notion of Givenness has been shown to be crucial in explaining the data pattern in a series of experiments on the comprehension of locative inversion constructions, since both effects on reading times found in these experiments (Relatum=Given, and the well-established Given-before-New effect) make reference to it (cf. Hörnig et al., 2005, and Hörnig et al., to appear; but s.also Clifton & Frazier (2004) for a slightly different view on the role of Givenness in comprehension).

It is on these grounds that I think arguments against the role of Givenness in the explanation of information structural partitions should be put under close scrutiny, and even if Givenness-based explanations of information structural phenomena were only to be had at a certain price, they may eventually just be worth it.

Bibliography

Clifton, Charles, & Lyn Frazier (2004). Should given information come before new? Yes and no. *Memory and Cognition*, 32(6):886–895.

Garrod, Simon & Anthony J. Sanford (1994). Resolving Sentences in a Discourse Context: How Discourse Representation Affects Language Understanding. In Morton Ann Gernsbacher, editor, *Handbook of Psycholinguistics*, pages 675–698. Academic Press, New York.

Hörnig, Robin, Klaus Oberauer & Andrea Weidenfeld (2005). Two principles of premise integration in spatial reasoning. *Memory & Cognition*, 33(1):131–139.

Hörnig, Robin, Thomas Weskott, Reinhold Kliegl & Gisbert Fanselow (to appear). Word order variation in spatial descriptions with adverbs. *Memory & Cognition*.

Jäger, Gerhard (2004). Alternatives or presuppositions? A comparison of the Background-Presupposition Rule with Alternative Semantics. *Theoretical Linguistics*, 30(1):111–122.

Kasimir, Elke (2005). Questions-Answers Test and Givenness: Some Question Marks. In Ishihara, Shinichiro, Michaela Schmitz & Anne Schwarz, editors, *ISIS, vol. 3*, pages 1–52. Universitätsverlag, Potsdam.

Krifka, Manfred (2004). The Semantics of Questions and the Focusation of Answers. In Lee, Chungmin, Matthew Gordon & Daniel Büring, editor, *Topic and Focus: Cross-linguistic Perspectives on Intonation and Meaning*, pages 139–151. Kluwer, Dordrecht.

Lepore, Ernest & Kirk Ludwig (2000). The Semantics and Pragmatics of Complex Demonstratives. *Mind*, 109.434:199–240.

Reinhart, Tanya (1981). Pragmatics and Linguistics: An Analysis of Sentence Topics. *Philosophica*, 27 (1):53–94.

Wagner, Michael (2005). Givenness and Locality. Handout for a talk held at University of Potsdam, January 18, 2005.

Thomas Weskott
SFB 632, Project C1
Universität Potsdam
Postfach 601553
14415 Potsdam
Germany
weskott@uni-potsdam.de
www.sfb632.uni-potsdam.de/homes/weskott/

The Semantics of Ellipsis[*]

Paul Elbourne

Queen Mary, University of London

There are four phenomena that are particularly troublesome for theories of ellipsis: the existence of sloppy readings when the relevant pronouns cannot possibly be bound; an ellipsis being resolved in such a way that an ellipsis site in the antecedent is not understood in the way it was there; an ellipsis site drawing material from two or more separate antecedents; and ellipsis with no linguistic antecedent. These cases are accounted for by means of a new theory that involves copying syntactically incomplete antecedent material and an analysis of silent VPs and NPs that makes them into higher order definite descriptions that can be bound into.

Keywords: VP-ellipsis, NP-deletion, definite descriptions

1 Introduction

There is a common view of ellipsis according to which an elided phrase[1] requires a linguistic antecedent (Hankamer and Sag 1976) and the relationship between elided phrase and antecedent is one of identity of Logical Form (LF) or meaning (Sag 1976, Williams 1977).

Hankamer and Sag (1976) argued that ellipsis had to have a linguistic antecedent on that basis of examples like these:

---

[1]I use the term *elided phrase* simply as a descriptive term, without wishing to advocate the view on which such phrases are underlyingly present and deleted in the phonology. The same goes, later, for my use of the term *NP-deletion*.

(1)   *(Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop.)*
         Sag: It's not clear that you'll be able to.

(2)   *(Sag produces a cleaver and makes as if to hack off his left hand.)*
         Hankamer: Don't be alarmed. He never actually does.

We are supposed to imagine these examples being acted out, as it were, so that there is no linguistic antecedent for the ellipsis in each case, only an accompanying action. Hankamer and Sag claim that (1) cannot felicitously be understood as "It's not clear that you'll be able to push that ball through that hoop," even though it is obvious what action is being referred to. An analogous claim is made for (2). Hence the requirement for linguistic antecedents.

Ellipsis is thought to be based on the meaning or LF of the antecedent phrase because of cases like the following, which is taken from Heim and Kratzer 1998. In (3), the elided VP can only be understood in the same way as the scopally ambiguous antecedent VP. That is, if we understand the antecedent with *a drawing* scoping above *every teacher*, we have to understand the elided VP this way too; and analogously for the other scopal construal.

(3)   Laura showed a drawing to every teacher, but Lena didn't.

If the resolution of ellipsis makes reference directly to the meaning of the antecedent and requires us to understand the same meaning at the ellipsis site, it is obvious that we can achieve this result. We can also achieve this result by supposing that ellipsis resolution makes reference to the syntactic level of LF, where the quantifiers will have moved to positions that reflect their scopal ordering in the semantics (May 1977, 1985). For example, we might have something like (4) as an LF representation of (3).

(4)   Laura T $[_{VP}$ a drawing$_1$ $[_{VP}$ every teacher$_2$ $[_{VP}$ show t$_1$ to t$_2$]]]
         Lena did not $[_{VP}$ a drawing$_1$ $[_{VP}$ every teacher$_2$ $[_{VP}$ show t$_1$ to t$_2$]]]

Then we could suppose that ellipsis consists simply of copying an antecedent

LF into the ellipsis site, or of optionally not pronouncing a piece of syntactic structure when its LF is identical to that of some antecedent.

This view is commonly linked to an account of strict and sloppy readings that sees them as deriving from the pronouns in the antecedent being referential and bound respectively. The strict and sloppy readings of a representative example are given in (5) (Ross 1967).

(5)   a.   John loves his mother and Bill does too.

       b.   '...Bill loves Bill's mother.' (Sloppy)

       c.   '...Bill loves John's mother.' (Strict)

If the pronouns in the VPs are ambiguous between referential and bound, we have a neat account of this ambiguity (Keenan 1971). We can suppose that the sloppy reading results from the pronoun being bound, as in (6), and the strict reading results from the pronoun being referential, as in (7), where it is to be understood that the sentence is interpreted with respect to a variable assignment that maps 1 to John.

(6)   *Sloppy*
       John [λ_2 t_2 love his$_2$ mother] and Bill [λ_2 t_2 love his$_2$ mother]

(7)   *Strict*
       John [λ_2 t_2 loves his$_1$ mother] and Bill [λ_2 t_2 love his$_1$ mother]

Again, reference to either the meaning or the LF of the antecedent would suffice for the correct interpretation to be obtained at the ellipsis site.

The view that the strict-sloppy ambiguity is to be dealt with in terms of referential versus bound pronouns is independent in principle of the theses that ellipsis requires a linguistic antecedent and that the relationship between antecedent and elided phrase is one of identity of meaning or LF; but in practise these three theories are often combined. I will call the composite view that assumes all three *the common view*.

It has been known for some time that the common view is at best only partially enlightening and that it may very well be utterly false. There are four phenomena that seem to be incompatible with it.

The first phenomenon is that of sloppy readings appearing when they cannot possibly be the result of pronouns in VPs being bound. Some examples, with sources, follow:

(8)   The policeman who arrested John read him his rights, but the policeman who arrested Bill didn't. (Wescoat 1989)

(9)   John's coach thinks he has a chance, and Bill's coach does too. (Rooth 1992)

(10)   If John has trouble at school, I'll help him, but if Bill does, I won't. (Hardt 1999)

(11)   *(John and Bill both have cats.)* When I met John, I talked to his cat, but when I met Bill, I didn't.

Take (8), for example. It clearly has a reading, "...but the policeman who arrested Bill didn't read Bill his rights." But this cannot be the result of the pronouns in the antecedent VP being bound. If they were bound, the antecedent would have the denotation $[\lambda x.\, x$ read x x's rights]. If one understands this after the subject of the second sentence, one obtains the meaning "The policeman who arrested Bill didn't read himself his own rights." This is clearly not the meaning that the sentence in fact has. Analogous considerations hold for sentences (9)–(11). Some though not all speakers obtain a reading for (9) that can be paraphrased "John's coach thinks John has a chance and Bill's coach thinks Bill has a chance." If I say (10), I say that if Bill has trouble at school I will not help him, Bill; and if I say (11) I say that when I met Bill I did not talk to his, Bill's, cat. None of these examples can be accounted for by the theory that sees sloppy readings of pronouns as arising from VP-internal pronouns being bound. They have in common the feature that the intuitive antecedent of the

relevant pronoun in the antecedent VP does not c-command it. In the sentences above, for example, *John* cannot c-command the pronouns *him, his* or *he* in the first conjunct. I will call readings like these *binderless sloppy readings*.

The second phenomenon that seems incompatible with the common view arises in connection with sentences like the following:

(12)   When John had to cook, he didn't want to. When he had to clean, he didn't either. (Hardt 1999, Schwarz 2000)

This example clearly has a reading "When John had to cook, he did not want to cook, and when he had to clean, he did not want to clean." How it arrives at this reading is entirely mysterious on the common view. The ellipsis in the first sentence seems straightforward enough. We take *cook* to be the antecedent, and resolve the ellipsis so as to produce a meaning "When John had to cook, he did not want to cook." The VP of the matrix clause in the first sentence will be [$_{\text{VP}}$ want to cook]; or if the ellipsis is not resolved in the syntax but at some level of semantic representation, there must be a VP denotation something like [$\lambda x.\, x$ wants to cook]. This VP is the only plausible antecedent for the ellipsis in the matrix clause of the second sentence. But any resolution procedure reliant on identity of meaning or LF structure then predicts that the second sentence will have to mean "When John had to clean, he did not want to cook." This is not the case, however. Thus the common view faces another significant problem. I will call examples like these examples of *ellipsis-containing antecedents*.

Note that the problem of ellipsis-containing antecedents arises in other configurations than that just given, where the antecedent for VP-ellipsis contained VP-ellipsis. The following examples involve NP-deletion:

(13)   Every police officer who arrested some murderers insulted some, and every police officer who arrested some burglars did too. (Elbourne 2001)

(14)   After the books went on sale, thirteen shoppers who had bought some earlier complained; but after the magazines went on sale, only two did.

(Eytan Zweig, personal communication)

(15)   When John wanted to cook, he met some people who didn't want him to; and when he wanted to clean he met some too.

(13), on one natural reading, means "Every police officer who arrested some murderers insulted some murderers and every police officer who arrested some burglars insulted some burglars."[2] There is NP-deletion in the first conjunct: *insulted some* is understood as "insulted some murderers." We then have VP-ellipsis in the second conjunct: *did too* intuitively takes *insulted some* as its antecedent; but instead of being understood as "insulted some murderers," it is understood as "insulted some burglars." An exactly analogous problem arises in connection with (14), which means "After the books went on sale, thirteen shoppers who had bought some books earlier complained; but after the magazines went on sale, only two shoppers who had bought some magazines earlier complained." So the problem arises also when the antecedent of NP-deletion contains NP-deletion. The fourth logical possibility is VP-ellipsis within the antecedent of NP-deletion, and we see this in (15). On one reading, this means "When John wanted to cook, he met some people who didn't want him to cook; and when he wanted to clean he met some people who didn't want him to clean." Again, there is no obvious way in which the common view, which posits straightforward identity of meaning or LF structure between antecedent and ellipsis, can account for these examples.

The third problem that faces the common view arises when an ellipsis site seems to be related to more than one antecedent, and to draw material from both. Some well-known examples are the following:

(16)   Bob wants to sail round the world and Alice wants to climb Kilimanjaro, but neither of them can, because money is too tight. (Webber 1978)

---

[2]There is possibly an ambiguity between "insulted some of the murderers he arrested" and "insulted some other murderers." This is not relevant here. See Elbourne 2001 for further discussion.

(17)   I did everything Mary did. Mary swam the English Channel and Mary climbed Kilimanjaro, and I did too. (Fiengo and May 1994)

(18)   Whenever Max uses the fax or Oscar uses the Xerox, I can't. (Fiengo and May 1994)

The interpretations of these examples are tricky. (17) is the easiest. It pretty clearly means "...and I swam the English Channel and climbed Kilimanjaro too." One wants to paraphrase (18) "Whenever Max uses the fax or Oscar uses the Xerox I cannot use the fax or the Xerox," but attention must be paid that we do not analyze the sentence as meaning "...I cannot do either." The correct interpretation seems to be something like "...I cannot use whichever one is being used." Similarly, the ellipsis in (16) cannot be resolved "...neither of them can do either," but must mean something like "neither of them can do the thing they want." These facts are clearly beyond the ability of the common view to capture. There are similar cases that involve NP-deletion, as pointed out in Elbourne 2001:

(19)   John needs a hammer. Mary needs a mallet. They're going to borrow Bill's.

This seems to mean "They're going to borrow Bill's hammer and mallet." I will refer to examples like these as involving *split antecedents*.

The fourth problem for the common view is that some cases of VP-ellipsis and NP-deletion require no linguistic antecedents whatsoever. Many people find the original examples of Hankamer and Sag (1976), repeated here, quite felicitous:

(20)   *(Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop.)*
       Sag: It's not clear that you'll be able to.

(21)   *(Sag produces a cleaver and makes as if to hack off his left hand.)*
       Hankamer: Don't be alarmed. He never actually does.

And there are many examples about which there is no debate whatsoever. The following list of VP-ellipsis cases is adapted from Merchant forthcoming, which should be consulted for original references:

(22)   *(John attempts to kiss Mary while driving.)*
        John, you mustn't.

(23)   *(A piece of chocolate cake is offered.)*
        I really shouldn't.

(24)   *(As an invitation to dance.)*
        Shall we?

(25)   *(Mary gets John an expensive present.)*
        Mary, you shouldn't have!

(26)   *(Gesturing towards an empty chair.)*
        May I?

(27)   *(Responding to the last.)*
        Please do.

(28)   *(Seeing someone about to do a shot of Tequila.)*
        If you can, I can too.

(29)   *(Looking at someone psyching herself up to jump across a wide gap.)*
        I bet she won't.

(30)   *(Seeing someone who has just died his hair green.)*
        You didn't!

(31)   *(Sitting next to someone doing something annoying.)*
        Must you?

(32)   *(On witnessing someone about to do anything undesirable.)*
        Don't!

Once again there are analogous cases of NP-deletion:

(33)   *(There are lots of barking dogs in the yard. We look at them without speaking. I point and say:)*

      Harry's is particularly noisy. (Elbourne 2001)

I will call this phenomenon *ellipsis with no linguistic antecedent*.

We have four problems, then, that the common view does not seem able to deal with, those of binderless sloppy readings, ellipsis-containing antecedents, split antecedents and ellipsis with no linguistic antecedent. In this article I lay out a theory of ellipsis that is compatible with all these data (section 2). I comment on relevant previous literature in section 3. Section 4 concludes.

It should be emphasized that I will be concentrating entirely on the semantics of ellipsis, without paying attention to any syntactic constraints there may be on when ellipsis is possible. (See Johnson 2001 for a good overview of syntactic and semantic issues connected with VP-ellipsis.) And I will be concentrating on VP-ellipsis and NP-deletion, without attempting to account for sluicing, pseudo-gapping and other kinds of ellipsis.

2 A Theory of Ellipsis

2.1   The Framework

2.1.1   Event Semantics and Little v

Following much work in event semantics and argument structure, I assume that VPs are predicates of events (Davidson 1967, Parsons 1990, Tenny and Pustejovsky 2000, Pylkkänen 2002), and that subjects are introduced by a special head v that takes the VP as its sister (Kratzer 1996, Pylkkänen 2002).

To give a flavor of the general approach, let us take the example in (34), which will have the structure in (35). The lambda-abstractor in the syntax is from Heim and Kratzer 1998.

(34)   Brutus stabbed Caesar.

(35)   [$_{\text{TP}}$ Brutus [λ_2 [T$_{\text{past}}$ [$_{\text{vP}}$ t$_2$ [v [$_{\text{VP}}$ stab Caesar]]]]]]

The following lexical entries will enable vPs and VPs to be predicates of events, with the whole sentence an assertion of the existence of certain kinds of time intervals and events.[3]

(36) $[\![\text{T}_{\text{past}}]\!] = \lambda p_{\langle \text{s,t} \rangle}.\exists t(t < \text{NOW} \ \& \ \text{at } t : \exists e \, p(e) = 1)$

$[\![\text{v}]\!] = \lambda p_{\langle \text{s,t} \rangle}.\lambda y.\lambda e.p(e) = 1 \ \& \ \text{Agent}(e, y)$

$[\![\text{stab}]\!] = \lambda z.\lambda e.\text{stabbing}(e) \ \& \ \text{Theme}(e, z)$

These lexical entries give the meanings displayed in Figure 1 for different parts of the tree. The sentence is interpreted with respect to a variable assignment g. Figure 1 is not necessarily a serious contender for what the syntax and semantics of this sentence actually look like. But it will be useful to have something concrete to work with.

2.1.2   Pronouns and Names

I will follow Postal (1966), Stockwell, Schachter and Partee (1973), Abney (1987), Longobardi (1994), Uriagereka (1995), Elbourne (2001, forthcoming) and Neale (2005) in assuming that pronouns are basically determiners. In particular, third person pronouns are definite articles whose complements must be phonologically null (Elbourne 2001, forthcoming); these complements can be referential indices, which I take to be of type $\langle \text{e,t} \rangle$, or they can be normal NPs unpronounced because of NP-deletion. For example, the index 2 in what we would normally write as *he*$_2$ might be interpreted, by means of a variable assignment mapping 2 to John, as $[\lambda x.x = \text{John}]$; since *he* has the same meaning

---

[3]Some head should presumably convert the denotation of the whole sentence into a set of possible worlds or situations, and there might also be heads that contribute illocutionary force. I omit all these for simplicity's sake and write as if the denotations of sentences were truth values. I also do not properly take account of the indexical nature of tense, which presumably must make reference to the time of utterance. I gesture towards this with the term NOW in the metalanguage, which is supposed to be an indexical taking as its value the time of utterance on each occasion of use.

$\boxed{\textbf{TP,}\ \exists t(t < \text{NOW} \ \& \ \text{at } t : \exists e(\text{stabbing}(e) \ \& \ \text{Theme}(e, \text{Caesar}) \ \& \ \text{Agent}(e, \text{Brutus})))}$

Brutus $\boxed{\lambda x.\exists t(t < \text{NOW} \ \& \ \text{at } t : \exists e(\text{stabbing}(e) \ \& \ \text{Theme}(e, \text{Caesar}) \ \& \ \text{Agent}(e, x)))}$

λ_2 $\boxed{\textbf{T}',\ \exists t(t < \text{NOW} \ \& \ \text{at } t : \exists e(\text{stabbing}(e) \ \& \ \text{Theme}(e, \text{Caesar}) \ \& \ \text{Agent}(e, g(2))))}$

$\textbf{T}_{\textbf{past}}$ $\boxed{\textbf{vP,}\ \lambda e.\text{stabbing}(e) \ \& \ \text{Theme}(e, \text{Caesar}) \ \& \ \text{Agent}(e, g(2))}$

$\textbf{t}_2$ $\boxed{\lambda y.\lambda e.\text{stabbing}(e) \ \& \ \text{Theme}(e, \text{Caesar}) \ \& \ \text{Agent}(e, y)}$

$\textbf{v}$ $\boxed{\textbf{VP,}\ \lambda e.\text{stabbing}(e) \ \& \ \text{Theme}(e, \text{Caesar})}$

stab Caesar
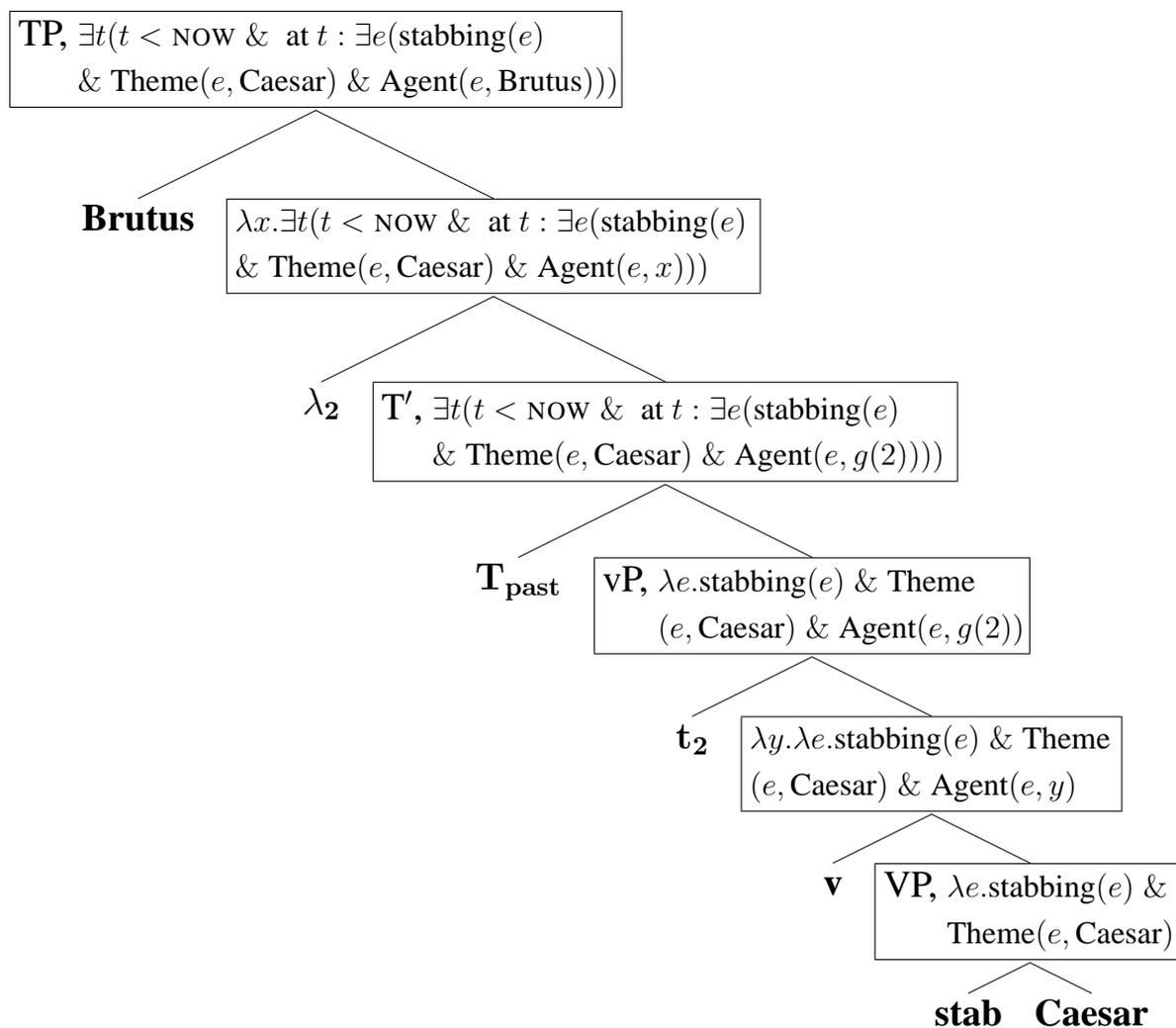
Figure 1: Brutus stabbed Caesar

as *the* (give or take ϕ-features), the interpretation of pronoun plus index in this case will be "the unique x such that x is identical to John," or, in other words, "John." This position has the advantage of unifying the referential and bound occurrences of pronouns with their use as donkey anaphors (Elbourne forthcoming), assuming a theory whereby donkey pronouns are analyzed as definite descriptions (Cooper 1979, Neale 1990, Heim 1990, Elbourne 2001, forthcoming).

I will follow Burge (1973), Recanati (1993), Larson and Segal (1995) and

Elbourne (forthcoming) in assuming that names are basically nouns. We often see them occurring with overt determiners, as in (37).

(37)   a.   An embattled Tony Blair addressed the Commons this afternoon.

        b.   Which Alfred did you mean? This Alfred?

When they appear to stand alone, they will be preceded by a special phonologically null definite determiner THE. This is paralleled by those languages like Classical Greek and some dialects of German in which names are regularly preceded by an overt definite article.

As for the semantics of names on this view, Burge's (1973) basic idea is that, for example, *Alfred* means something like "entity called Alfred," and variants of this have been proposed by the other authors just cited. In Elbourne forthcoming I propose that on most occasions of use *Alfred* will mean "entity called Alfred and identical to a," where a is an individual constant picking out a particular entity called Alfred. In this article I will just assume things like $[\lambda x.x$ is an Alfred] for the meaning of names, since their exact semantics is orthogonal to the issues of primary concern.

I will just assume that nouns are of type $\langle e,t \rangle$, and that definite articles, including pronouns, are functions from predicates of type $\langle e,t \rangle$ to individuals (Heim 1991, Heim and Kratzer 1998, von Fintel 2004, Elbourne 2005, forthcoming), as proposed originally by Frege (1893). The semantics for some relevant lexical items is shown in (38).[4]

(38) $[\![\text{the}]\!]^{g,h} = \lambda f_{\langle e,t \rangle}.\iota x\, f(x) = 1$

        $[\![\text{him}]\!]^{g,h} = \lambda f_{\langle e,t \rangle}.\iota x\, f(x) = 1$

---

[4]The semantics given in (38) is a simplification in that we probably need to embed the whole lexical entry in a situation semantics or possible worlds semantics for full adequacy, and have definite articles be functions from properties to individual concepts, i.e. functions from circumstances of evaluation to individuals, as in Elbourne 2005, forthcoming. I overlook this complication here and continue to operate with an extensional semantics. I also overlook the ϕ-features on the pronouns, for the sake of simplicity.

$[\![\text{cat}]\!]^{g,h} = \lambda x . x$ is a cat

$[\![\text{Alfred}]\!]^{g,h} = \lambda x . x$ is an Alfred

The semantics of the metalanguage operator $\ulcorner \iota \urcorner$ is as follows: for any function f, the denotation of $\ulcorner \iota x\, f(x) = 1 \urcorner$ will be of type e, if it is defined; if there is exactly one entity x such $f(x) = 1$, the denotation of $\ulcorner \iota x\, f(x) = 1 \urcorner$ will be that very individual; if there is no such individual, the whole expression will have no value. (So the expression in effect introduces a presupposition that there is exactly one such individual, since an utterance containing it will not be felicitous otherwise.) The individual that is the value of the expression will naturally vary from model to model. For example, if our universe is $\{2, 3, 4\}$, then the denotation of $\ulcorner \iota x\, x > 3 \urcorner$ is 4; if the universe is $\{2, 3, 5\}$, the value of the same expression will be 5. This, simply put, is how definite descriptions differ from constants.

2.1.3   Ellipsis

In this section I will sketch a theory of ellipsis that will enable us to give a straightforward account of the sentences involving ellipsis-containing antecedents and binderless sloppy readings, and I will apply it to some relatively simple data. In the next sections I will use to analyze the problematic data that we saw in section 1.

The theory is as follows:

(39)   *Theory the First*

VP-ellipsis and NP-deletion consist in the generation of bare VP and NP nodes, respectively. These structures are sent to PF. There is an LF process of resolving the ellipsis, whereby the bare nodes are replaced with a copy of a phrase of the same syntactic category drawn from the linguistic environment.

$$\begin{array}{c}
\text{TP} \\
\diagdown \\
\text{DP} \qquad \text{T}' \\
\end{array}$$
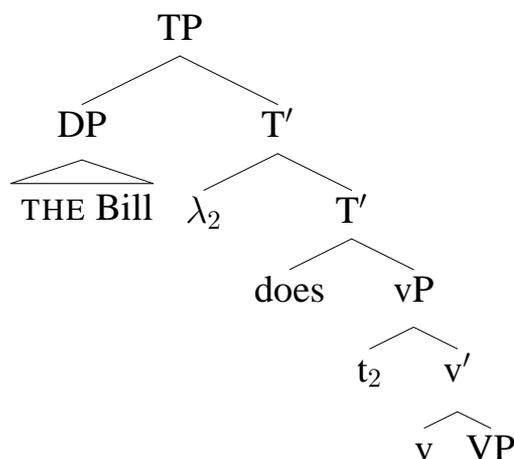
Figure 2: Bill does too

According to this account, then, a sentence involving VP-ellipsis or NP-deletion begins life as a structure that is syntactically incomplete. For example, the last sentence of (40) will have the (possibly simplified) structure in Figure 2.

(40)   John loves Mary and Bill does too.

Note that in Figure 2 we have a VP node that is simply not spelled out any further. This will be possible if we adhere to a traditional conception of phrase structure rules that allows things like (41).

(41)   v′ → v VP

It is obviously incompatible with Chomsky's (1995) Bare Phrase Structure, according to which the idea of a phrasal node with no daughters does not make sense. I hope to show that significant empirical advantages can be gained from the traditional conception of syntactic rules.

We generate, then, a structure like that in Figure 2, and this is what is pronounced. Ellipsis resolution will then be an LF process that replaces the bare phrasal node with a copy of a phrase of the same syntactic category drawn from the linguistic environment.[5] In the case of the current example, then, we copy the antecedent VP [VP love Mary] and replace the empty VP node with it.

---

[5]We will need to make an addition to our theory when we return to the consideration of

Before we go on to look at how this conception of ellipsis facilitates the analysis of ellipsis-containing antecedents and binderless sloppy readings, there are a couple of potential problems to address. First, we should pause to consider the case of simple bound variables in VP-ellipsis. Consider (42).

(42)   Every woman loves her mother. Even Mary does.

According to almost all current theories, *her* in the first sentence of this example is to be analyzed as being or containing a bound individual variable.[6] I too will assume this. The current theory, then, must maintain that we start off with the LF structure in (43).

(43)   Every woman λ_2 T t_2 v love her$_2$ mother. Even THE Mary λ_2 does t_2 v
        VP.

The process of LF ellipsis resolution produces the following:

(44)   Every woman λ_2 T t_2 v love her$_2$ mother. Even THE Mary λ_2 does t_2 v
        love her$_2$ mother.

This produces the right reading, of course. And the use of the same index on the expressions bound by the subjects of the first and second sentences violates no prohibition that I know of. Difficulties with repeated use of the same index only arise if the same index is used on referential expressions with different intended referents, or on bound variables intended to be bound by different operators that lie within the scope of both (unlike in this case), or on both bound variables and an independent referential expression. Heim and Kratzer (1998:254) have proposed a principle explicitly to deal with the latter case.

---

ellipsis with no linguistic antecedent in section 2.5. Meanwhile, if the idea of replacing a node with something else causes unease, one could also think of ellipsis resolution as copying the daughters of a node of the same category in the linguistic environment and pasting them separately into position beneath the ellipsis node. But I personally find the version in the text less awkward.

[6]An exception is the variable-free semantics proposed by Szabolcsi (1989) and explored by Jacobson (1999, 2000) in connection with Categorial Grammar. I will not attempt to assess this work here. See Elbourne forthcoming for some critical discussion.

The second matter that a simple copying theory of ellipsis, like the present one, must address, is what Fiengo and May (1994: 218) call *vehicle change*. The question is to deal with examples like the following:

(45)   I turned in my assignment, but most of the other students didn't.

The problem, of course, is that a straightforward theory of copying (or, indeed, of deletion under identity) seems to predict that the last sentence here will only be able to mean "most of the other students didn't turn in *my* assignment," when in fact it can mean that most of the other students did not turn in *their* assignment. This need not be seen as ruling out the current approach, however. One theory that has been proposed is to see pronouns like *my* here as simple bound variables semantically devoid of ϕ-features, their ϕ-features being inherited from their binders by an agreement process at PF (Kratzer 1998, Rullmann 2004, Heim 2005), and for present purposes I will assume that something like this is the case.

2.2   Ellipsis-Containing Antecedents

Let us reconsider an example of an ellipsis-containing antecedent, (12), repeated here as (46).

(46)   When John had to cook, he didn't want to. When he had to clean, he didn't either.

Recall that the problem is that (46) can mean "... when he had to clean, he didn't want to clean."

It can be seen that with the theory of ellipsis just sketched, the problem is resolved quite easily. We start out with a simplified LF representation like that in (47) for the utterance in question.

(47)   when John had to cook, he did not want to v VP
        when he had to clean, he did not v VP

There are bare VP nodes here, and we have a choice which one we fill in first. Suppose we take the second and replace it with a copy of the matrix VP in the first sentence. We obtain the following:

(48)   when John had to cook, he did not want to v VP

        when he had to clean, he did not v want to v VP

We can then fill in the resultant bare VP nodes with simple VPs drawn from the respective preceding sentences:

(49)   when John had to cook, he did not want to v cook

        when he had to clean, he did not v want to v clean

The right meaning results. The other cases of ellipsis-containing antecedents, (13)–(15), will work analogously.

We might very well wonder if this system overgenerates. The simple answer is that it does indeed, if we suppose it not to be supplemented with other considerations. For example, the theory as it stands predicts that the following will also be a possible LF structure for (46):

(50)   when John had to cook, he did not want to v clean

        when he had to clean, he did not v want to v cook

There is nothing in Theory the First to prevent us reaching the stage shown in (48) and then looking forward to the second sentence and filling in the ellipsis in the first with the VP *clean*, and looking back to the first sentence and understanding *cook* in the final ellipsis site. But I take it that these ellipsis resolutions will be ruled out by independent factors. After all, the syntactic structure in (48) is exactly that which we see overtly spelled out in (51):

(51)   When John had to cook, he didn't want to. When he had to clean, he

        didn't want to either.

And it is a fact that this example cannot be understood as in (50) either. I take it then that Theory the First is correct as far as it goes, but that it must be sup-

plemented with other considerations, perhaps having to do with processing, that restrict the VPs that can be used as antecedents. Perhaps distance from the ellipsis site is one heuristic: if we consider (48) as a linear string, we see that the VP *clean* is six VPs back from the ellipsis site for which it is being considered as a possible antecedent.

2.3   Binderless Sloppy Readings

I will now consider how we should deal with cases of binderless sloppy readings. Let us reconsider (10), repeated here as (52):

(52)   If John has trouble at school, I'll help him, but if Bill does, I won't.

Recall that this example can mean, "...if Bill has trouble at school I won't help Bill."

Given the analysis of pronouns as definite articles and names as predicates outlined in section 2.1.2, this case in fact reduces to the last one. Binderless sloppy readings, in other words, are also cases of ellipsis-containing antecedents. The initial (slightly simplified) structure of the current example is the following:

(53)   if THE John T v have trouble at school, I will v help him NP
        if THE Bill does v VP, I will not v VP

Note that it makes sense under the current conception of pronouns to say that there is NP-deletion after *him*, since *him* is a determiner. (It would differ from the normal definite article *the* in allowing NP-deletion after it.) Hence the bare NP node following *him* in (53). The resolution of the first VP-ellipsis is straightforward:

(54)   if THE John T v have trouble at school, I will v help him NP
        if THE Bill does v have trouble at school, I will not v VP

We now replace the remaining bare VP node with a copy of the matrix VP of the first sentence:

(55)   if THE John T v have trouble at school, I will v help him NP
         if THE Bill does v have trouble at school, I will not v help him NP

And we resolve the two instances of NP-deletion in the most straightforward way by taking antecedents in the respective sentences of the bare NP nodes:

(56)   if THE John T v have trouble at school, I will v help him John
         if THE Bill does v have trouble at school, I will not v help him Bill

Again, the correct meaning is obtained. The other examples of binderless sloppy readings will work analogously.

2.4   Split Antecedents

Recall the examples of split antecedents in (16)–(18), repeated here as (57)–(59).

(57)   Bob wants to sail round the world and Alice wants to climb Kilimanjaro, but neither of them can, because money is too tight.

(58)   I did everything Mary did. Mary swam the English Channel and Mary climbed Kilimanjaro, and I did too.

(59)   Whenever Max uses the fax or Oscar uses the Xerox, I can't.

(57) seems to be interpreted something like "...neither of them can do what they want to do." (58) means "...I swam the English Channel and climbed Kilimanjaro too." And (59) seems to mean "...I cannot use whichever machine is being used."

Given that these interpretations do not appear to have very close syntactic links to any antecedent Verb Phrases, it is tempting at this point to say that we have been on the wrong track all along, and that the interpretation of an elided

VP can be any property of events that the hearer might reasonably be expected to work out. But this would be going too far. Consider (60), which is taken from Heim 1996.

(60)   The garbage can is full. *I hope that *you* will, for a change.

It is obvious here that the speaker means "I hope that you will take out the garbage." But despite the fact that it is easy to work out the intended meaning, this example does not work as a VP-ellipsis. There must be a tighter connection with some previous VP.

One rarely considered option that might nevertheless be explored at this stage is to say that the elided VP can be interpreted as any property of events that has *some* syntactic connection with an antecedent VP. For example, we can interpret (57) as "...neither of them can do what they want to do" because we have the word *want* in a preceding VP. On this theory, (60) would not felicitous because no plausible VP meaning can be reconstructed that would use any functions contributed by any word in the antecedent VP *is full*. In particular, neither of these words contributes anything from which the meaning "take out the garbage" can be constructed. This is too unimaginative, though. If we are allowed to create "...neither of them can do what they want to do" solely on the basis of the word *want* and a shrewd idea of what the speaker might be driving at, then surely we could construct "make the garbage can not be full any more" from the word *full* and the same kind of shrewd idea. But, to repeat, (60) does not seem to be a successful VP-ellipsis, no matter what precise way we think of understanding it.

I conclude that in these cases, then, we still have a very close connection to the antecedents. In particular, I assume that reference must be made to the exact form of the antecedent VPs, as in Theory the First.[7] As a first step, let us reexamine the examples and see if we can come up with paraphrases that seem

---

[7]We will return to the cases of ellipsis with no linguistic antecedent in section 2.5.

to incorporate the exact meanings of the antecedents. My proposal is that our examples are to be paraphrased as in the (b) sentences below, where the phrases in italics hark back to the antecedent VP denotations.

(61)   a.   Bob wants to sail round the world and Alice wants to climb Kilimanjaro, but neither of them can, because money is too tight.

       b.   Bob wants to sail round the world and Alice wants to climb Kilimanjaro, but neither of them can perform the particular action or actions out of *sailing round the world* and *climbing Kilimanjaro* that they desire.

(62)   a.   I did everything Mary did. Mary swam the English Channel and Mary climbed Kilimanjaro, and I did too.

       b.   Mary swam the English Channel and Mary climbed Kilimanjaro and I performed the particular action or actions out of *swimming the English Channel* and *climbing Kilimanjaro* that Mary performed.

(63)   a.   Whenever Max uses the fax or Oscar uses the Xerox, I can't.

       b.   Whenever Max uses the fax or Oscar uses the Xerox, I can't perform the particular action or actions out of *using the fax* and *using the Xerox* that are being performed.

It can be seen that one form of paraphrase covers all the examples. Informally, in place of the elided VP we understand "perform the particular action or actions out of f_1 and f_2 that have property F," for VP meanings f_1 and f_2 and properties of VP meanings F.

I propose to spell out parts of the above paraphrase schema with LF operators. For example, the LF of the final sentence of (62a) will be that shown in Figure 3. There is a special set of lexical items with the following semantics:

(64)   For all $n > 0$, $[\![\text{AND}^n]\!]^g = \lambda f_{1,\langle s,t\rangle} \ldots f_{n,\langle s,t\rangle}.\lambda h_{\langle s,t\rangle}.h \leq_i f_1 \oplus \ldots \oplus f_n$

The notation is that of Link's (1983) theory of plurality. An operator AND^n takes n arguments of type $\langle s,t\rangle$ and maps them to the characteristic function of the set

```
                                        TP
                                      /    \
                                   I_5      T'
                                          /    \
                                       λ_4      T'
                                              /    \
                                         T_past    vP
                                                  /    \
                                                t_4     v'
                                                      /    \
                                                     v      THEP
                                                          /        \
                                                      THE'          AND^0P
                                                     /    \        /        \
                                                  THE     RP    AND^1P        VP
                                                          |    /      \       /    \
                                                    R_1,⟨st,t⟩ AND^2   VP  climb Kilimanjaro
                                                                     /    \
                                                            swim the English Channel
```
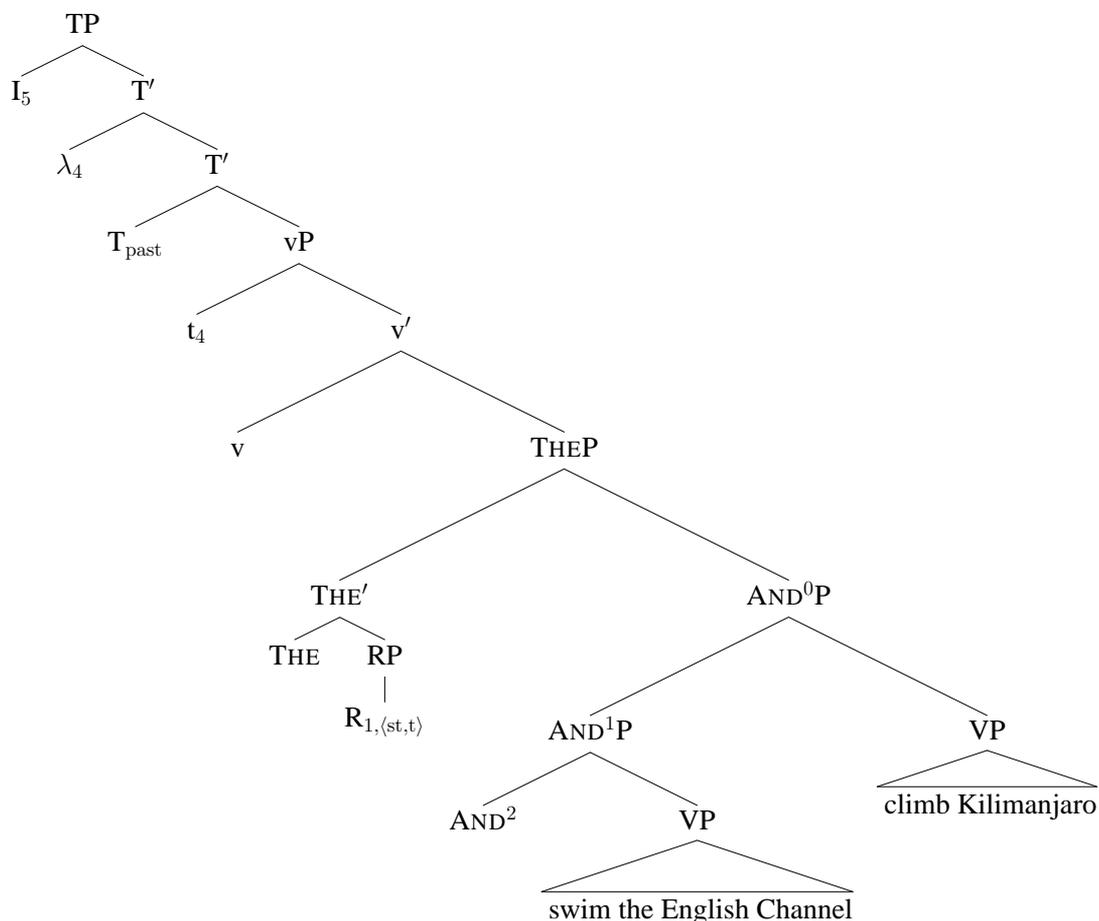
Figure 3: I did too...

of $\langle s,t \rangle$ functions that are part of the plural individual that has all and only the n arguments as atomic parts. In the present case, we have the following:

(65) $[\![\text{AND}^2]\!]^g = \lambda f_{1,\langle s,t \rangle}.\lambda f_{2,\langle s,t \rangle}.\lambda h_{\langle s,t \rangle}.h \leq_i f_1 \oplus f_2$

This means that the denotation of AND^0P in the syntax is as in (66). I use the two italicized phrases to stand for the meanings of the two VPs.

(66) $\lambda h_{\langle s,t \rangle}.h \leq_i \textit{swimming} \oplus \textit{climbing}$

The point of THE and its argument $R_{1,\langle st,t \rangle}$ is to introduce the modification of the VP-meanings that we have seen to be necessary in some of the paraphrases in (61)–(63). In the present case, as it happens, this item is redundant, but I will show the argument $R_{1,\langle st,t \rangle}$ in action for the sake of illustration. (It will play a

central role in analysis of (61a) and (63a); it so happens that (62a) is simple in ways that make it a good introductory example.) Let us assume, then, that $R_{1,\langle st,t \rangle}$ is assigned the value shown in (67):

(67) $[\![R_{1,\langle st,t \rangle}]\!]^g = \lambda f_{\langle s,t \rangle}.\exists e(f(e) = 1 \ \& \ \text{Agent}(e, \text{Mary}))$

This function is not the value of any overt linguistic constituent, but we can assume that this does not matter for LF variables. The mention of Mary doing things makes this function salient enough.

Meanwhile, the operator THE has the denotation in (68), which uses some terminology from Link 1983 defined in (69); $\ulcorner *P \urcorner$ is the plural predicate, the one that characterizes both singular entities that are P and plural entities whose atomic parts are all P.

(68) $[\![\text{THE}]\!]^g = \lambda F_{\langle st,t \rangle}.\lambda G_{\langle st,t \rangle}.\sigma f(F(f) = 1 \ \& \ G(f) = 1)$

(69) $\sigma x Px := \iota x(*Px \ \& \ \forall y(*Py \rightarrow y \leq_i x))$

In other words, THE takes as its arguments two properties of VP-meanings and maps them to the maximal plural individual composed of individuals that satisfy the two arguments. (I use *individual* here not to mean an entity of type e but to mean an atom within the relevant domain, which is here $D_{\langle s,t \rangle}$.)

Given these definitions, the denotation of THEP in Figure 3 is (70a), which in the present context is equivalent to (70b).

(70)   a. $\sigma f(\exists e(f(e) = 1 \ \& \ \text{Agent}(e, \text{Mary})) \ \& \ f \leq_i swimming \oplus climbing)$

　　　　 b. $swimming \oplus climbing$

Moving upwards in Figure 3, we come to v and T_{past}, whose denotations we wrote in (36) as follows:

(71) $[\![T_{\text{past}}]\!]^g = \lambda f_{\langle s,t \rangle}.\exists t(t < \text{NOW} \ \& \ \text{at } t : \exists e \, f(e) = 1)$

　　　　 $[\![v]\!]^g = \lambda f_{\langle s,t \rangle}.\lambda y.\lambda e.f(e) = 1 \ \& \ \text{Agent}(e, y)$

These lexical entries will still suffice, but we now have to be sure to understand the notion of Agent in such a way that one can be an Agent of plural events. Let

us say that one is an Agent of a plural event if and only if one is an Agent, in the normal sense, of all the events that are atomic parts of it. We will also need to make sense of the notion of an event (a plural event, to be sure) satisfying a plural individual made up of VP-meanings. Let us say that for any event e and functions f, g of type $\langle \text{s,t} \rangle$, $f \oplus g(e) = 1$ if and only if there exist events e' and e'' such that $f(e') = 1$ and $g(e'') = 1$ and $e' \leq_i e$ and $e'' \leq_i e$.

Assuming the speaker is John, we finally arrive at the truth conditions in (72) for the whole sentence:

(72) $\exists t(t < \text{NOW} \ \& \ \text{at } t : \exists e(swimming \oplus climbing(e) = 1 \ \&$
        $\text{Agent}(e, \text{John})))$

In other words, there was in the past a plural event e such that e had as its parts an event of swimming the English Channel and an event of climbing Kilimanjaro and John was the agent of e, in the new sense whereby he was the agent of every atomic part of e. These truth conditions seem to be intuitively adequate.

I will shortly go on to analyze (61a) (the sentence about the globe-trotting desires of Bob and Alice), but before doing so I should perhaps be more explicit about the new syntax of VP-ellipsis than I have been so far. The proposal is that a vP can be spelled out by the rules and rule-schemas in (73):

(73)   vP $\rightarrow$ v THEP
        THEP $\rightarrow$ THE$'$ AND^{0}P
        AND^{n}P $\rightarrow$ AND^{n+1}P VP
        AND^{n}P $\rightarrow$ AND^{n+1} VP
        THE$'$ $\rightarrow$ THE RP
        RP $\rightarrow$ R$_{m,\langle \text{st,t} \rangle}$
        RP $\rightarrow$ R$_{m,\langle \text{e,stt} \rangle}$ pro$_{l,\text{e}}$

I am not aware of any cases where the RP has to contain more than one variable of type e, so I have just listed two cases above; a more sophisticated treatment along the lines of that given to AND^{0}P could be devised if necessary.
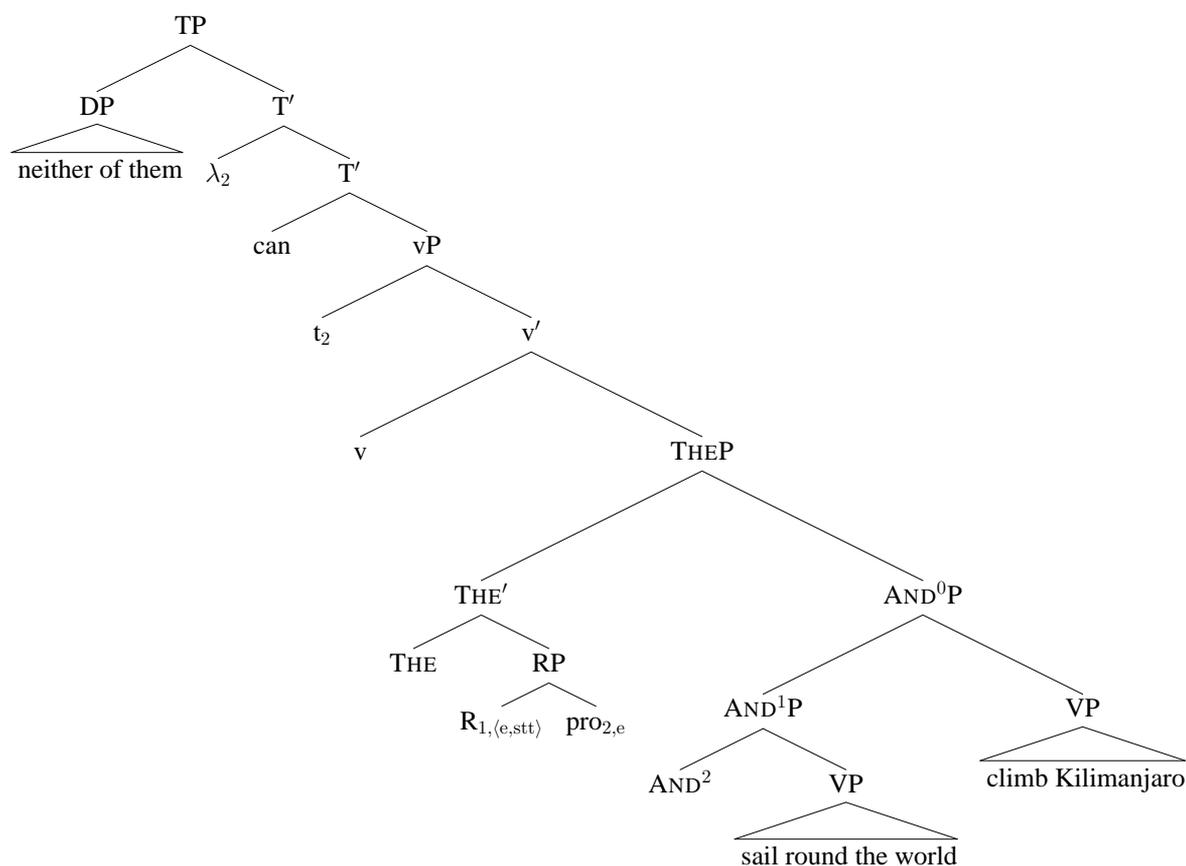
Figure 4: Neither of them can. . .

Let us move on to the analysis of (61a), repeated here as (74).

(74)   Bob wants to sail round the world and Alice wants to climb Kilimanjaro, but neither of them can, because money is too tight.

The LF for *neither of them can* will be that shown in Figure 4. The free variable $R_{1,\langle e,stt\rangle}$ will be assigned a meaning as follows:

(75) $[\![R_{1,\langle e,stt\rangle}]\!]^g = \lambda x.\lambda f_{\langle s,t\rangle}.x$ desires that there be an event e such that $f(e) = 1$ and $\text{Agent}(e, x))$

I will avail myself of the following simple denotations for *can* and *neither of them*:

(76) $[\![\text{can}]\!]^g = \lambda f_{\langle s,t\rangle}.$it is possible that there be an event e such that $f(e) = 1$
        $[\![\text{neither of them}]\!]^g = \lambda f_{\langle e,t\rangle}.\neg\exists x((x = \text{Bob} \vee x = \text{Alice})\ \&\ f(x) = 1)$

Given these denotations, the truth conditions for this example come out to be as in (77). I use italicized expressions to abbreviate meanings of the VPs.

(77) $\neg \exists x((x = \text{Bob} \vee x = \text{Alice})$ and it is possible that there be an event e' such that $\sigma f(x$ desires that there be an event e such that $f(e) = 1$ and $\text{Agent}(e, x)$ and $f \leq_i$ *sailing* $\oplus$ *climbing*$)(e') = 1$ and $\text{Agent}(e', x))$

In other words, there does not exist an individual x such that x is Bob or Alice and it is possible that x be the agent of an event that satisfies the unique predicate f such that x wants to be the agent of an f-event and f is one of *sailing round the world* and *climbing Kilimanjaro*. This seems to be intuitively adequate.[8]

(63a), repeated here as (78a), will work by the same means, as suggested by the paraphrase in (78b).

(78) a. Whenever Max uses the fax or Oscar uses the Xerox, I can't.

b. Whenever Max uses the fax or Oscar uses the Xerox, I can't perform the particular action or actions out of *using the fax* and *using the Xerox* that are being performed.

In other words, there will be two small VPs *using the fax* and *using the Xerox*, and an R variable will be assigned a denotation something like "currently being performed." Working out an exact analysis would require us to make decisions regarding what entities *whenever* quantifies over (time intervals? situations?) and whether these are represented in the syntax. The general outlines are clear,

---

[8]The idea of having the variable $R_{1, \langle e, stt \rangle}$ provide extra descriptive material to modify syntactically more robust material is reminiscent of the approach to quantifier domain restriction that posits variables in the syntax, as proposed by von Fintel 1994, Stanley 2000 and Stanley and Szabó 2000. In particular, von Fintel (1994) sometimes has two variables in such positions, one an individual variable bound by the subject, in order to deal with sentences like *Only one class was so bad that no student passed*, where we are to understand "only one class x...no student in x...." The combination of a definite article plus a relation variable plus an individual variable is also reminiscent of the LF configuration posited by Heim and Kratzer (1998) to spell out donkey pronouns.

however.[9]

There are also examples of split antecedents involving NP-deletion, as we saw in (19), repeated here as (79). I will analyze the variant in (80), which is more revealing of structure since the quantifier *each* actually seems to bind into the NP-deletion site.

(79)   John needs a hammer. Mary needs a mallet. They're going to borrow Bill's.

(80)   John needed a hammer. Mary needed a mallet. Each borrowed Bill's.

We will need a set of rules for spelling out silent NPs parallel to the ones we saw for VPs in (73). The rules and rule-schemas in (81) will suffice.

(81)   DP $\rightarrow$ D THEP
        THEP $\rightarrow$ THE$'$ AND^{0}P
        AND^{n}P $\rightarrow$ AND^{n+1}P NP
        AND^{n}P $\rightarrow$ AND^{n+1} NP
        THE$'$ $\rightarrow$ THE SP
        SP $\rightarrow$ S$_{m,\langle et,t \rangle}$
        SP $\rightarrow$ S$_{m,\langle e,ett \rangle}$ pro$_{l,e}$

Translating the proposal just explored with respect to VP-ellipsis into the NP domain, we arrive, then, at the slightly simplified LF in Figure 5 for the last sentence of (80). I ignore any complexity there may be behind the surface forms *each* and *Bill's*. The new operators THE and AND^2 will receive the interpretations in (82) and (83), parallel to the interpretations of THE and AND^2.

(82) $[\![\text{THE}]\!]^g = \lambda F_{\langle et,t \rangle}.\lambda G_{\langle et,t \rangle}.\sigma f(F(f) = 1 \,\&\, G(f) = 1)$

---

[9]The present apparatus can also be put into service to analyze cases of overt conjunction of VPs, as in *John walked and sang*. All that is needed is for the overt *and* in this position to mean $\lambda f_{\langle s,t \rangle}.\lambda g_{\langle s,t \rangle}.f \oplus g$. There are arguably conceptual advantages to having *and* produce a sum of entities when it appears between VPs, just as it does when it conjoins expressions of type e. See Krifka 1990 and Lasersohn 1995 for detailed proposals concerning the non-boolean conjunction of VPs, and Winter 2001 for discussion.

```
                              TP
                    ┌─────────┴─────────┐
                   DP                   T′
                    │          ┌────────┴────────┐
                  each        λ₂                 T′
                                        ┌────────┴────────┐
                                     T_past              vP
                                               ┌──────────┴──────────┐
                                              t₂                     v′
                                                           ┌─────────┴─────────┐
                                                           v                   VP
                                                    ┌──────┴──────┐
                                                 borrow           DP
                                                          ┌───────┴───────┐
                                                          D              THEP
                                                          │        ┌──────┴──────┐
                                                        Bill's    THE′         AND⁰P
                                                             ┌─────┴─────┐  ┌────┴────┐
                                                            THE         SP AND¹P     NP
                                                                   ┌────┴───┐ ┌──┴──┐  △
                                                              S₁,⟨e,ett⟩ pro₂,ₑ AND² NP mallet
                                                                                    △
                                                                                 hammer
```
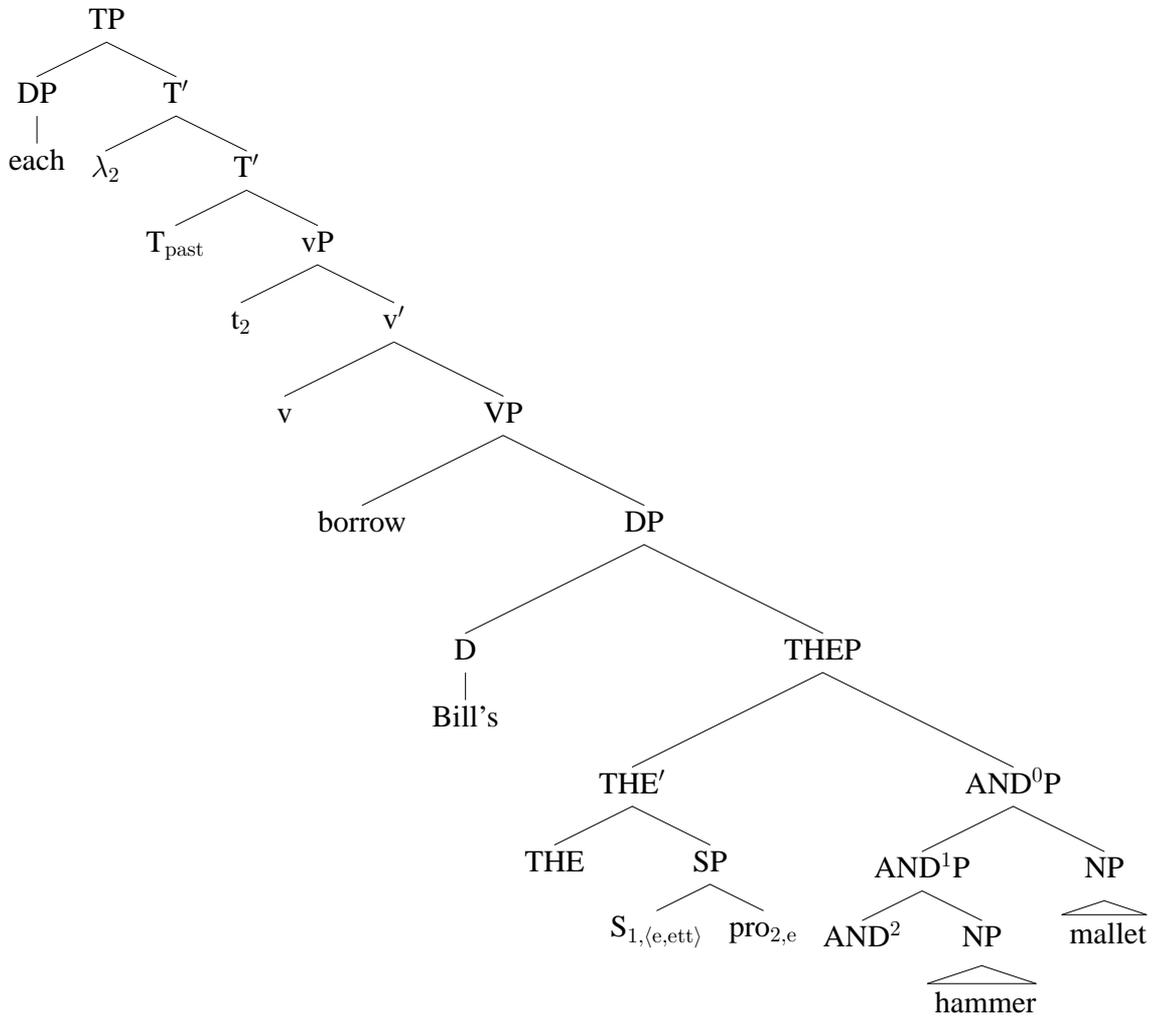
Figure 5: Each borrowed Bill's

(83) $[\![\mathrm{AND}^2]\!]^g = \lambda f_{\langle \mathrm{e,t}\rangle}.\lambda g_{\langle \mathrm{e,t}\rangle}.\lambda h_{\langle \mathrm{e,t}\rangle}.h \leq_i f \oplus g$

The nouns *hammer* and *mallet* receive the denotations one might expect, and the free variable $S_{1,\langle \mathrm{e,ett}\rangle}$ will receive the following interpretation from the variable assignment g:

(84) $[\![S_{1,\langle \mathrm{e,ett}\rangle}]\!]^g = \lambda x.\lambda f_{\langle \mathrm{e,t}\rangle}.\, x$ needs an f

(85) $[\![\mathrm{hammer}]\!]^g = \lambda x.x$ is a hammer

(86) $[\![\mathrm{mallet}]\!]^g = \lambda x.x$ is a mallet

I abstract away from the complexities inherent in the analysis of transitive intensional verbs like *need*. Allowing ourselves the convenient lexical entries in (87), (88) and (89) for *each*, *Bill's* and *borrow*, we arrive at the truth conditions in (90) for the last sentence of (80). I use italicized words to abbreviate the meanings of *hammer* and *mallet*.

(87) $[\![\text{each}]\!]^g = \lambda f_{\langle e,t \rangle}.\forall x((x = \text{John} \vee x = \text{Mary}) \rightarrow f(x) = 1)$

(88) $[\![\text{Bill's}]\!]^g = \lambda f_{\langle e,t \rangle}.\iota x(x \text{ is Bill's } \& f(x) = 1)$

(89) $[\![\text{borrow}]\!]^g = \lambda x.\lambda e.\text{borrowing}(e) \ \& \ \text{Theme}(e, x)$

(90) $\forall x((x = \text{John} \vee x = \text{Mary}) \rightarrow \exists t(t < \text{NOW} \ \& \ \text{at } t : \exists e(\text{borrowing}(e)$
$\& \ \text{Agent}(e, x) \ \& \ \text{Theme}(e, \iota y(y \text{ is Bill's } \& \ \sigma f(x \text{ needs an } f \ \& \ f \leq_i$
$hammer \oplus mallet)(y) = 1)))))$

The claim, then, is that the last sentence of (80) is true if and only if, for all x such that x is Mary or John, x was the Agent of a borrowing event whose Theme was the unique item of Bill's that satisfied the unique predicate f such that x needed an f and f was one of *hammer* and *mallet*. This seems to be accurate.

It is time to consider how to integrate the model that we have built up for split antecedent cases with the theory that we developed in previous sections for binderless sloppy readings and ellipsis-containing antecedents. Recall Theory the First in (39), repeated here as (91):

(91) *Theory the First*
VP-ellipsis and NP-deletion consist in the generation of bare VP and NP nodes, respectively. These structures are sent to PF. There is an LF process of resolving the ellipsis, whereby the bare nodes are replaced with a copy of a phrase of the same syntactic category drawn from the linguistic environment.

We can combine this theory with the procedures we have posited to deal with the split antecedent cases by adopting the following statement:

(92)   *Theory the Second*

    (i)   VP and NP nodes may be bare.

    (ii)   vPs may be spelled out as in (73).

    (iii)   DPs may be spelled out as in (81).

    (iv)   VPs and NPs in AND^0Ps and AND^0Ps must be bare.

    (v)   A bare VP or NP node must be replaced at LF by a copy of a phrase of the same syntactic category drawn from the linguistic environment.

In other words we view the trees in Figures 3–5 not as being base-generated but as deriving from structures that originally had bare VPs and NP nodes in their AND^0Ps and AND^0Ps. All else proceeds as previously described, and we still retain the option of handling the ellipsis-containing antecedent cases and binderless sloppy readings with the simpler structures posited earlier.[10]

Before we leave these data, we should note that it is also possible to concoct labored but not ungrammatical examples that combine the traits of the various species that we have been examining. (93), for example, is a combination of a split antecedent case and a case of an ellipsis-containing antecedent:

---

[10]An alternative way of unifying Theory the First with the structures posited for split antecedent cases would maintain that VP-ellipsis and NP-deletion always involve definite description structures of the type posited in (73) and (81). In cases without split antecedents we would just have one VP or NP as part of the structure, and the operator AND^1 or AND^1. If we suppose that some trivial property is generally available for the denotation of RP and SP when these phrases are redundant, it turns out that a definite description structure with just one NP or VP is semantically equivalent to just having the NP or VP there by itself. I marginally prefer the option given in the text because of the complexity of the structures that result in cases of ellipsis-containing antecedents if we suppose that we always have definite descriptions in ellipsis. But the issue is a subtle one, and the theory described in this note has a certain kind of unity that cannot be claimed by the one in the text.

(93)   When Bob had to sail round the world and Mary had to climb Kili-
        manjaro, they didn't want to; and when Bob had to swim the English
        Channel and Mary had to climb K2 they didn't, either.

It is simple to derive this example in the current theory: the overt *want to* is
followed by a little v and a silent THEP containing two bare VP nodes, and
before these VP nodes are filled in this structure is copied and used to resolve
the VP-ellipsis in the second sentence. There are then two separate processes of
resolving split antecedent ellipsis, one in each sentence.

2.5   Ellipsis with No Linguistic Antecedent

Recall the cases of ellipsis with no linguistic antecedent in (22)–(33), some of
which are repeated in (94)–(98).

(94)   *(John attempts to kiss Mary while driving.)*
        John, you mustn't.

(95)   *(A piece of chocolate cake is offered.)*
        I really shouldn't.

(96)   *(As an invitation to dance.)*
        Shall we?

(97)   *(Mary gets John an expensive present.)*
        Mary, you shouldn't have!

(98)   *(There are lots of barking dogs in the yard. We look at them without
        speaking. I point and say:)*
        Harry's is particularly noisy.

The question is how to integrate these cases into the framework developed in
section 2.4.

    Roughly speaking, what the examples of ellipsis with no linguistic antecedent
have in common is that there is some obvious sensory (in these cases visual)

clue to the property conveyed by the unpronounced phrase. The clue need not always be an instance of the relevant action or entity before the very eyes of the speaker and hearer: there is kissing in the scenario of (94) and a dog in that of (98), but not necessarily any eating in that of (95), and certainly not any eating of the piece of cake being offered; and there may or may not be dancing actually taking place when (96) is uttered, provided that it is clear to speaker and hearer that they are at a dance. But in the cases of VP-ellipsis there must at least be an obvious result of the action in question or a stimulus towards performing it. It is hard to be more precise, and I will leave the matter here for now, pending further research.

I propose that in these cases too we have full syntactic VPs and NPs at LF.[11] So in (94), for example, we might have something like [VP kiss me now]. We can now emend our theory to the following, which is the final version:

(99)   *Theory the Third*

   (i)   VP and NP nodes may be bare.

   (ii)   vPs may be spelled out as in (73).

   (iii) DPs may be spelled out as in (81).

   (iv)   VPs and NPs in AND⁰Ps and AND⁰Ps must be bare.

   (v)   A bare VP or NP node must be filled in at LF by a VP or NP that is highly salient.

   (v)   A VP or NP is highly salient if and only if:

---

[11]One sometimes gets the impression that some theorists think that verb-phrase meanings that are merely contextually salient or able to be worked out, as opposed to occurring as the value of some constituent in the linguistic environment, should not come to be represented as syntactically fully-fledged VPs. It is unclear what the grounds for this view could be, however. If it ever happens that we think of things and then put our thoughts into words, which is not implausible, we are extraordinarily adept at moving from non-linguistic to linguistic modes of representation.

(a)   its denotation describes an action or thing made salient by an obvious sensory clue; or

(b)   it is a copy of a phrase of the same syntactic category drawn from the linguistic environment.

This theory should be seen as part of general linguistic competence, but it will be used by speakers and hearers in different ways. The hearer, whose job it is to try to work out what the speaker was saying, might not arrive at the precise unpronounced LF phrase that is present in the mind of the speaker, especially in cases where there is no linguistic antecedent; but communication will have proceeded well enough if the speaker comes up with something with the same or a relevantly similar meaning. See Neale 2005 for salutary discussion of the asymmetric roles of speaker and hearer.

Before leaving the current theory, there are two loose ends that should be tied up. Firstly, the theory of ellipsis outlined here is naturally one way of spelling out some details left obscure in my previous work (Elbourne 2001, forthcoming) in which I claimed that E-type anaphora was NP-deletion; but this latter thesis is independent of any particular analysis of NP-deletion.

Secondly, we should revisit example (60), repeated here as (100), to make sure that we do not end up predicting that it is good.

(100)   The garbage can is full. *I hope that *you* will, for a change.

In the absence of a strong sensory clue of the kind exemplified earlier, we have to work with the antecedent VP *is full*. The example presumably fails because there is no sufficiently salient relation R, a possible first argument of THE, that could combine with the meaning of this VP and give the desired interpretation. Note that if a strong visual clue is offered, the example becomes better. If, for example, I hand you the brimming garbage can and utter the last sentence of (100), the sentence dramatically improves.

3 Previous Literature

In this section I briefly compare the theory advocated in this article with some other theories that try to cover some or all of the tricky cases dealt with here. I will not attempt a detailed review of the literature on ellipsis, which would be a mammoth undertaking.[12]

3.1   Rooth 1992

Rooth's influential paper proposes using entailment-like relations involving focus to characterize the relationship between elided phrase and antecedent. He claims that VP-ellipsis is permitted only if two conditions are met: first, the lexical content of the elided VP at LF must be the same as that of an antecedent VP, modulo indices on pronouns and traces; and second, the elided VP must be embedded in a constituent β such that there is a constituent α containing the antecedent VP such that the ordinary semantic value of α is a member of the set of focus alternatives generated by β. A variant of Rooth's condition is to be found in work by Merchant (2001).

Some of the details of Rooth's theory are less than satisfactory, however. He gives an analysis of binderless sloppy readings like (101) that crucially involves the DP *John* scoping out of its containing DP. This already seem dubious, since DP is normally an island. And we surely cannot extend this idea to (102) and

---

[12]Two theories worthy of note that I do not deal with in the main text are the higher order unification theory of Dalrymple, Shieber and Pereira (1991) and the discourse grammar theory of Prüst, Scha and van den Berg (1994). The former deals with binderless sloppy readings and might be extended to deal with (62a), the sentence about Mary swimming the English Channel and climbing Kilimanjaro; but it does not attempt to deal with ellipsis-containing antecedents and the other split antecedent cases, and I see no way of extending it so that it would. Prüst, Scha and van den Berg (1994) also account for cases like (62a) but they do not attempt to deal with the other split antecedent cases, the binderless sloppy readings or the ellipsis-containing antecedents, and again I see no way of extending their theory to achieve better coverage.

(103), where the relevant DP would have to scope out of a relative clause or an *if*-clause.

(101)   John's coach thinks he has a chance, and Bill's coach does too.

(102)   The policeman who arrested John read him his rights, but the policeman who arrested Bill didn't.

(103)   If John has trouble at school, I'll help him, but if Bill does, I won't.

Rooth does not offer any analysis of ellipsis-containing antecedents or split antecedents.

3.2   Fiengo and May 1994

Fiengo and May (1994) attempt to account for binderless sloppy readings (and other data) by means of a complex system that exploits isomorphism of patterns of indices in trees. I will not attempt to summarize this theory here.[13]

Fiengo and May also analyze split antecedent cases such as (104).

(104)   I did everything that Mary did. Mary swam the English Channel, and Mary climbed Kilimanjaro, and I did too.

About this example, they say the following (1994:195):

> In this sentence, what is elided are occurrences of the VPs *swim the English Channel* and *climb Kilimanjaro* (and an occurrence of *and*). This is all we need to know to "recover" the ellipsis—that

---

[13]We should note, however, that Rooth already in his 1992, page 18, had published the following counterexample to Fiengo and May's theory:

  (i)   Yesterday John's boss told him to shape up, and today Bill's boss did.

  (ii) Yesterday the guy John works for told him to shape up, and today Bill's boss did.

Since *John* and *Bill* do not occupy isomorphic positions in their respective sentences in (ii), it is unclear that Fiengo and May's account correctly predicts the sloppy reading here.

> is, that the final clause is *I swam the English Channel and climbed Kilimanjaro*. That the elided occurrences must be conjoined in the representation of ([104]) just follows from the way they can "fit" into its structure.

They further suggest the following (1994:200), referring to examples like (78a) ("Whenever Max uses the fax or Oscar uses the Xerox…"):

> As a general rule, the discourse sentence is also the domain from which the elided coordinating element is drawn. In a case of disjunction, for instance, *or* is reconstructed.

It appears, then, that Fiengo and May are assuming that a conjunction must be reconstructed somehow on the basis of the linguistic environment. A possible problem with this theory is that there are split antecedent cases when there is no *and* or *or* in the linguistic environment (Elbourne 2001):

(105)   Mary swam the English Channel. Mary climbed Kilimanjaro. I did too.

It is unclear to me how (105) could be dealt with in Fiengo and May's theory of split antecedent cases. It poses no problem for the theory advocated in this article, of course, since this theory does not rely on a conjunction being present in the linguistic environment.

Fiengo and May do not deal with cases of ellipsis-containing antecedents.

3.3   Hardt 1999

As far as I know, Hardt's (1999) theory is the only one previously published that attempts to account for binderless sloppy readings, ellipsis-containing antecedents and split antecedents. It is also not difficult to see how it might be extended to deal with cases of ellipsis with no linguistic antecedent. For the purpose of illustration, I will give an informal summary of how Hardt analyzes binderless sloppy readings and ellipsis-containing antecedents.

Hardt uses a dynamic semantics incorporating a notion of *discourse center*, based on the Centering framework of Grosz, Joshi and Weinstein 1995. In particular, discourse representation structures (the "boxes" of traditional DRT (Kamp 1981)) contain special discourse markers (variables) assigned to the *center* (roughly, topic) of the discourse. This enables Hardt to explain binderless sloppy readings along the following lines. Take (106).

(106)   If Tom was having trouble in school, I would help him. If Harry was having trouble, I wouldn't.

In the first sentence, the discourse center is Tom. The overt pronoun *him* is translated by a special discourse marker assigned to the center. So *help him* means something like "help the current center." This meaning is then understood at the ellipsis site; more precisely, it is the assigned as the value of the INFL of the second sentence, which is a deictic element in this theory. In the meantime, however, the center has changed. Harry is the discourse center of the second sentence. So the second sentence ends up meaning "I wouldn't help Harry," as desired.

The same principle can be used in cases of ellipsis-containing antecedents. Take our standard example:

(107)   When John had to cook, he didn't want to. When he had to clean, he didn't either.

Hardt supposes that in the first sentence the discourse center is the property of cooking. The first VP-ellipsis is resolved by having *to* be a deictic expression that picks up the property that is the current center. We arrive at the meaning "want to cook," then, for the end of the first sentence. The value of the deictic INFL of the second sentence is taken to be the DRT representation of *want to*, something interpretable as "want to perform the kind of action that is the current center." And again by the time we get to this point the center has shifted,

according to Hardt. It is now the property of cleaning, and the so the correct interpretation is obtained.

Even from this informal summary, it can be seen that it is important to Hardt's system that there is no syntactic structure, or at least no syntactic complexity, at ellipsis sites. The correct interpretation is arrived at by assigning denotations involving variables picking out the current center to the INFL of the ellipsis sentence. (One could also imagine a variant in which there was a deictic element in the VP position, as opposed to having INFL do this job.) But this feature of the theory, which lends it a useful flexibility and power, means that it is ill-equipped to deal with cases where there seems to be movement from ellipsis sites, as in the following examples.

(108)   Which book did John read? And which book did Bill?

(109)   John read every book that Bill did.

See Johnson 2001 for a summary of the controversy on whether theories without normal syntactic structures in the ellipsis sites can deal with examples like these. The upshot is not encouraging for those theories, and things seem especially difficult for the particular version that Hardt puts forward, according to which there is nothing whatsoever in ellipsis sites. By contrast, the theory advocated in the current article has normal syntactic structure in all ellipsis sites.[14]

---

[14]Hardt's theory also faces a knotty technical difficulty in analyzing certain seemingly simple cases of sloppy identity. Take *Tom loves his cat and John does too* (Hardt 1999:194). Using the device described above, Hardt analyzes *loves his cat* as roughly "loves the cat of the current center," and wants this meaning understood at the ellipsis site, by which time the center has changed to John. But the actual DRT representation used to express this meaning by Hardt also, necessarily given the system, contains a discourse marker for Tom's cat. Roughly speaking, and in slightly incongruous terms, we can think of *his cat* here as meaning something like "the unique x such that x is the cat of the current center and x is identical to c," where c is a constant referring to Tom's cat. If we understand this at the ellipsis site, then, we obtain a contradiction: the claim would be that John loves the cat of the current center (himself) that is identical to c (not his own, but Tom's cat). To avoid this difficulty Hardt proposes to reconstruct an "alphabetic variant" of the original property, one that replaces the troublesome discourse marker referring to Tom's cat by another one (Hardt 1999:195). But this seems like the merest

3.4   Tomioka 1999

Tomioka 1999 analyzes binderless sloppy readings by proposing that the pronouns in such cases are actually donkey pronouns. He follows Cooper (1979) in supposing that donkey pronouns are interpreted as definite descriptions containing bound variables. So the overt *read him his rights* in (110) means something like "read the person he arrested his rights," with "he" bound by the subject.

(110)   The policeman who arrested John read him his rights, but the policeman who arrested Bill didn't.

We understand the same meaning at the ellipsis site and the correct reading is obtained.

We know, however, that the descriptive content of donkey pronouns cannot be obtained so flexibly, just by picking up contextually salient relations. If it could, the following examples would have the same status (Heim 1990, Elbourne 2001):

(111)   a.   Every man who has a wife is sitting next to her.

        b.   ?? Every married man is sitting next to her.

In particular, (111b) would be fully felicitous, since *her* could mean "the person he is married to." This is not the case, however, meaning that the mechanism relied on by Tomioka is problematic. There are also problems with the assumption that donkey pronouns can contain bound individual variables. See Elbourne 2001 for further discussion.

Tomioka (1999) does not discuss ellipsis-containing antecedents or split antecedents.

---

stipulation. Alphabetic variants are not equivalent in dynamic systems, unlike in traditional logics. It matters whether we say, for example, $\exists x P x$ or $\exists y P y$, since the former but not the latter will be able to bind a syntactically free variable x that occurs in a later formula. As far as I can see, then, it is not only stipulative but actually illegal to solve the current problem by relying on the notion of alphabetic variance.

3.5   Schwarz 2000

Schwarz (2000, Chapter 5) analyzes cases of ellipsis-containing antecedents that involve VP-ellipsis within VP-ellipsis by having the VPs scope out and bind variables in both their overt positions and the ellipsis sites. So (112) has the LF in (113).

(112)   When John had to cook, he didn't want to. When he had to clean, he didn't either.

(113)   [cook] λQ[when John had to Q he didn't want to Q]
        [clean] λQ[when he had to Q he didn't want to Q]

It can be seen that the correct interpretation would result, and the antecedent and elided VPs are now identical. But surely we should assume, unless forced to do otherwise, that LF movement of VPs respects what we know about islands. Phrases cannot, of course, generally move out of *when*-clauses. The theory advocated in the current article does not posit any abnormal movement.

   Schwaz (2000) does not attempt to analyze binderless sloppy readings or split antecedents.

4 Conclusion

I will refrain from summarizing my theory in this section, since I have done so already in (99). I will merely note two possible extensions of it that could profitably be explored in future research. The question at issue concerns to what extent the new definite description structure for VPs and NPs, which I have cautiously posited so far only in split antecedent cases of ellipsis, should extended to other syntactic categories and other kinds of occasion.

   First, since I have concentrated on VP-ellipsis and NP-deletion in this article, we should ask whether the LF apparatus introduced here has counterparts in other kinds of ellipsis too, such as pseudo-gapping and sluicing. The answer

presumably depends on whether other kinds of ellipsis display split antecedent effects, since it was these cases that necessitated the new structures for NPs and VPs in the current theory. A possible indicator in the case of sluicing is (114):

(114)   Either John called someone or Mary called someone, but I don't know who.

This seems to have a reading "…but I don't know who was called by whichever one of them it was." This split antecedent interpretation constitutes evidence for extending the present theory at least to sluicing.

Second, we should ask whether some pronounced VPs and NPs might have the kind of definite description structure posited in this article. Evidence that this might be so comes from the so-called "respectively" readings of sentences like (115).

(115)   You and I did everything that Mary and Jane did. Mary swam the English Channel, and Jane climbed Kilimanjaro, and you and I swam the English Channel and climbed Kilimanjaro too.

Fiengo and May (1994:197) point out this sentence has a reading "You and I swam the English Channel and climbed Kilimanjaro respectively." In other words, we get the split antecedent interpretation without ellipsis in cases like this. Reverting to our former schema, we might paraphrase it "You and I performed the action or actions out of *swimming the English Channel* and *climbing Kilimanjaro* that were done by the person we were imitating." This is strong evidence for elements of the VP structure posited in this article being present in a pronounced VP.

Bibliography

Abney, Steven. 1987. *The English Noun Phrase in its sentential aspect.* Doctoral dissertation, MIT. Distributed by MITWPL.

Burge, Tyler. 1973. Reference and proper names. *The Journal of Philosophy* 70, 425–39.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Cooper, Robin. 1979. The interpretation of pronouns. In Frank Heny and Helmut Schnelle, eds., *Syntax and Semantics 10: Selections from the Third Gröningen Round Table,* 61–92. New York: Academic Press.

Dalrymple, Mary, Stuart Shieber and Fernando Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14: 399–452.

Davidson, Donald. 1967. The logical form of action sentences. In Nicolas Rescher, ed., *The logic of decision and action,* 81–95. Pittsburgh: University of Pittsburgh Press.

Elbourne, Paul. 2001. E-type anaphora as NP-deletion. *Natural Language Semantics* 9(3), 241–288.

Elbourne, Paul. 2005. Presuppositional demonstratives. Ms., Universität Potsdam.

Elbourne, Paul. Forthcoming. *Situations and Individuals.* Cambridge, Mass.: MIT Press.

Fiengo, Robert and Robert May. 1994. *Indices and Identity,* Cambridge, MA: MIT Press.

von Fintel, Kai. 1994. Restrictions on Quantifier Domains. Doctoral dissertation, University of Massachusetts, Amherst.

von Fintel, Kai. 2004. Would you believe it? The King of France is back! (Presuppositions and truth-value intuitions.) In Anne Bezuidenhout and

Marga Reimer, eds., *Descriptions and beyond: an interdisciplinary collection of essays on definite and indefinite descriptions and other related phenomena,* 315–341. Oxford: Clarendon Press.

Frege, Gottlob. 1893. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik* 100, 25–50. [Translated as 'On sense and reference', in Peter Geach and Max Black, eds., *Translations from the philosophical writings of Gottlob Frege,* 56–78. Oxford: Blackwell.]

Grosz, Barbara, Aravind Joshi and Scott Weinstein. 1995. Centering: a framework for modelling the local coherence of discourse. *Computational Linguistics* 21: 203–225.

Hankamer, Jorge and Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7, 391–426.

Hardt, Dan. 1999. Dynamic interpretation of verb phrase ellipsis. *Linguistics and Philosophy* 22, 187–221.

Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13, 137–177.

Heim, Irene. 1991. Artikel und Definitheit. In Arnim von Stechow and Dieter Wunderlich, eds., *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung,* 487–535. Berlin: Walter de Gruyter.

Heim, Irene. 1996. Lecture notes on ellipsis. Lecture notes, MIT.

Heim, Irene. 2005. Features on bound pronouns: semantics or syntax? Talk given at Semantiknetzwerk workshop, ZAS, May 2005.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in generative grammar.* Oxford: Blackwell.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22, 117–184.

Jacobson, Pauline. 2000. Paycheck pronouns, Bach-Peters sentences, and variable-free semantics. *Natural Language Semantics* 8, 77–155.

Johnson, Kyle. 2001. What VP ellipsis can do, and what it can't, but not why. In Mark Baltin and Chris Collins, eds., *The Handbook of Contemporary Syntactic Theory,* 439–479. Oxford: Blackwell.

Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, Theo Janssen and Martin Stokhof, eds., *Formal methods in the study of languages*, 277–322. Amsterdam: Mathematical Centre.

Keenan, Edward. 1971. Names, quantifiers, and a solution to the sloppy identity problem. *Papers in Linguistics* 4(2).

Kratzer, Angelika. 1996. Severing the external argument from its verb. In J. Rooryck and L. Zaring, eds., *Phrase Structure and the Lexicon,* 109–137. Dordrecht: Kluwer.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In *Proceedings from Semantics and Linguistic Theory VIII,* ed. by Devon Strolovitch and Aaron Lawson, 92–110. Ithaca, N.Y.,: Cornell University, CLC Publications.

Krifka, Manfred. 1990. Boolean and non-boolean *and*. In László Kálmán and László Pólos, eds., *Papers from the Second Symposium on Logic and Language.* Budapest: Akadémiai Kiadó.

Larson, Richard and Gabriel Segal. 1995. *Knowledge of Meaning. An Introduction to Semantic Theory.* Cambridge, MA: MIT Press.

Lasersohn, Peter. 1995. *Plurality, Conjunction and Events.* Dordrecht: Kluwer.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze and Arnim von Stechow, eds., *Meaning, use and interpretation of language,* 302–323. Berlin: Walter de Gruyter.[Also published in Godehard Link, *Al-*

gebraic semantics in language and philosophy, 11–34, Stanford: CSLI Publications, 1998.]

Longobardi, Giuseppe. 1994. Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25, 609–665.

May, Robert. 1977. The grammar of quantification. Doctoral dissertation, MIT, Cambridge, MA.

May, Robert. 1985. *Logical Form. Its Structure and Derivation.* Cambridge, MA: MIT Press.

Merchant, Jason. 2001. *The syntax of silence. Sluicing, islands, and the theory of ellipsis.* Oxford: Oxford University Press.

Merchant, Jason. Forthcoming. Fragments and ellipsis. *Linguistics and Philosophy.*

Neale, Stephen. 1990. *Descriptions.* Cambridge, MA: MIT Press.

Neale, Stephen. 2005. Pragmatism and binding. In Zoltán Gendler Szabó, ed., *Semantics versus pragmatics,* 165–285. Oxford: Clarendon Press.

Parsons, Terence. 1990. *Events in the semantics of English: a study in subatomic semantics.* Cambridge, MA: MIT Press.

Postal, Paul. 1966. On so-called 'pronouns' in English. In F. Dinneen, ed., *Report on the seventeenth annual round table meeting on linguistics and language studies,* 177–206. Washington, D.C.: Georgetown University Press.

Prüst, Hub, Remko Scha and Martin van den Berg. 1994. Discourse grammar and verb phrase anaphora. *Linguistics and Philosophy* 17: 261–327.

Pylkkänen, Liina. 2002. Introducing arguments. Doctoral dissertation, MIT.

Recanati, François. 1993. *Direct reference. From language to thought.* Oxford: Blackwell.

Rooth, Mats. 1992. Ellipsis redundancy and reduction redundancy. In Stephen Berman and Arild Hestvik, eds., *Proceedings of the Stuttgart Ellipsis Workshop.* Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29. Heidelberg: IBM Germany.

Ross, John. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, MA.

Rullmann, Hotze. 2004. First and second person pronouns as bound variables. *Linguistic Inquiry* 35: 159–168.

Sag, Ivan. 1976. Deletion and Logical Form. Doctoral dissertation, MIT.

Schwarz, Bernhard. 2000. *Topics in ellipsis.* PhD dissertation, University of Massachusetts at Amherst.

Stanley, Jason. 2000. Context and Logical Form. *Linguistics and Philosophy* 23: 391–434.

Stanley, Jason and Zoltan Szabó. 2000. On quantifier domain restriction. *Mind & Language* 15: 219–261.

Stockwell, Robert, Paul Schachter and Barbara Hall Partee. 1973. *The major syntactic structures of English.* New York: Holt, Rinehart and Winston.

Szabolcsi, Anna. 1989. Bound variables in syntax: are there any? In Renate Bartsch, Johan van Benthem and Peter van Emde Boas, eds., *Semantics and Contextual Expression,* 295–318. Dordrecht: Foris.

Tenny, Carol, and James Pustejovsky, eds. 2000. *Events as grammatical objects. The converging perspectives of lexical semantics and syntax.* Stanford: CSLI Publications.

Tomioka, Satoshi. 1999. A sloppy identity puzzle. *Natural Language Semantics* 7, 217–241.

Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26, 79–123.

Webber, Bonnie. 1978. A formal approach to discourse anaphora. Doctoral thesis, Harvard University.

Wescoat, Michael. 1989. Sloppy readings with embedded antecedents. Ms., Stanford University.

Williams, Edwin. 1977. Discourse and logical form. *Linguistic Inquiry* 8, 101–139.

Winter, Yoad. 2001. *Flexibility Principles in Boolean Semantics. The interpretation of coordination, plurality, and scope in natural language.* Cambridge, MA: MIT Press.

Paul Elbourne
Department of Linguistics
School of Modern Languages
Queen Mary, University of London
Mile End Road
London E1 4NS
United Kingdom
p.d.elbourne@qmul.ac.uk

Out-of-focus Encoding in Gur and Kwa[*]

Ines Fiedler and Anne Schwarz

Humboldt-Universität zu Berlin

This paper investigates the structural properties of morphosyntactically marked focus constructions, focussing on the often neglected non-focal sentence part in African tone languages. Based on new empirical evidence from five Gur and Kwa languages, we claim that these focus expressions have to be analysed as biclausal constructions even though they do not represent clefts containing restrictive relative clauses. First, we relativize the partly overgeneralized assumptions about structural correspondences between the out-of-focus part and relative clauses, and second, we show that our data do in fact support the hypothesis of a clause coordinating pattern as present in clause sequences in narration. It is argued that we deal with a non-accidental, systematic feature and that grammaticalization may conceal such basic narrative structures.

Keywords: ex-situ *focus, focus marker, relative clause, conjunction, grammaticalization*

1   Introduction

This paper deals with a phenomenon concerning marked *ex-situ* focus constructions which is known from several West African languages, among them Hausa, Fulfulde and others, namely the existence of structural parallels of the out-of-focus part of these constructions with relative (and other subordinated) clauses and partly also with narrative clause types. In Hausa for example, there are two morphosyntactical codings for the perfective and the

imperfective aspect: a canonical paradigm A and a paradigm B which is not only found in focus constructions, but also in relative clauses (henceforth RC) and – with respect to the perfective aspect – in narrative clauses (henceforth NC). This second marked paradigm is often called the "relative" form of the respective tense/aspect and its distribution has been discussed by some authors (cf. Bearth 1993, Frajzyngier 2004).

Not so well-known up to now is the fact that similar phenomena, comprising relative and/or narrative structures in pragmatically and linguistically marked sentences, do also appear in languages of the Gur and Kwa group studied by us. In these *ex-situ* focus constructions, a focused nominal constituent takes the sentence initial position.[1] We will present their structural features in relation to the language-specific relative and narrative clause types and discuss the implications of our findings from diachronic and comparative perspectives.

Our language sample consists of five Ghanaian languages which we have been investigating in the field in 2004. Its Gur part consists of the two languages Buli and Dagbani which belong to different subgroups of the Oti-Volta branch and which are spoken in the Northern area of Ghana. The three Kwa languages considered are also of different subgroups and comprise the Inland dialect of Ewe (Gbe), the Asante dialect of Akan (Potou Tano) and the Togo mountains language Lelemi (*na*-Togo). The status of the Togo remnant or Togo mountains languages as belonging to Gur or Kwa seems however still under discussion according to Rongier 1997 (cited in Blench 2001).

---

[1] For reasons of space, we will use the term focus constituent here also in those cases in which only part of the clause constituent is focal and we will discuss in this paper only affirmative constructions. In-situ focus phenomena are not considered here at all.
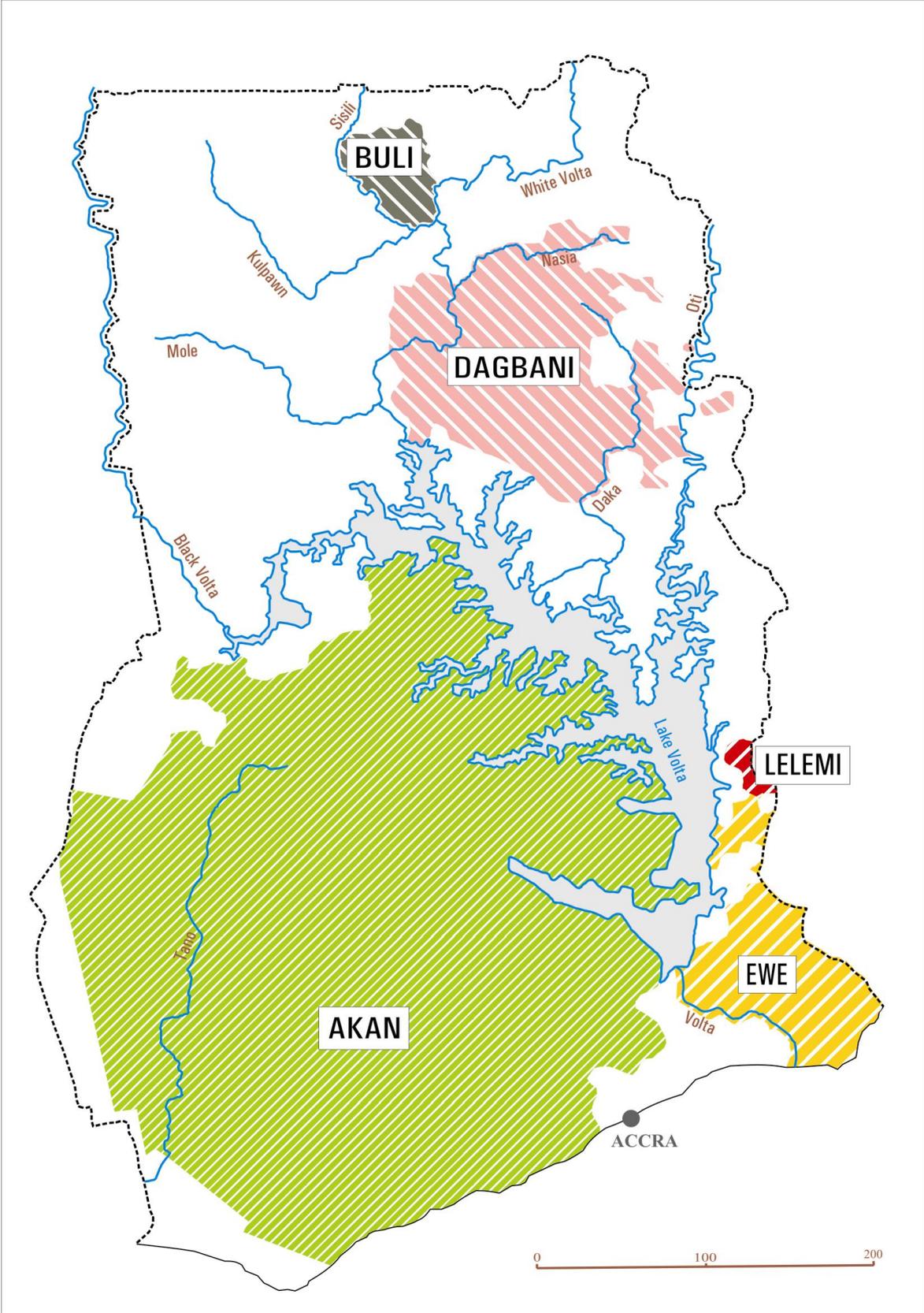
Illustration 1: Ghana Map showing our language sample

The structure of the paper is as follows: In part 2, we first present evidence for three structural characteristics that are recurring in the *ex-situ* focus constructions of almost all five languages, starting with the asymmetry between *ex-situ* subject und non-subject focus constructions (henceforth SF resp. NSF) (2.1), going on to relative-like features of these constructions (2.2) up to parallels with narrative clauses (2.3). In part 3, we first give an overview of the constructions' distribution (3.1) before we discuss the interpretation and the degree of grammaticalization of the narrative structures in each of the studied languages, and describe our findings comparatively in the conclusion (4.).

2   Structural Features of *ex-situ* Focus Constructions

In this chapter, we will demonstrate three observations concerning the structure of morphosyntactically marked focus constructions.[2]

2.1   SF vs. NSF Asymmetry

There is a constant structural asymmetry between SF and NSF constructions. This asymmetry shows up in several ways in the selected languages. Formally, we don't find the same degree of asymmetry in all the languages considered here. We will start with cases that are less obvious at first sight, and go on further to languages showing a full range of this asymmetry.

2.1.1 Ewe

The characteristics of *ex-situ* focus constructions in Ewe are as follows: First, the focused element can be marked by a FM (*y)é*, which is obligatory for subject focus and optional for non-subject focus. The exact constraints for the optionality are not yet clear. Second, there are special subject pronouns for 2nd

---

[2]   Please note that the data are transcribed with surface tones and that versalia in the English translation indicate the respective focal part of the utterance. In examples providing restrictive RCs, the relativized head and the RC are underlined.

and 3rd singular person which are used only in NSF constructions, while in SF no pronominals are preposed to the verb.

SF

(1)   ńtsú-á-é   tsɔ́-è.
        man-DEF-YE   take-o:3sg
        'The MAN took it.' (not the woman)

NSF

(2)   èdzì(-é)   wò-   ɖù.
        top(-YE)   3sg.DEP   eat
        'He was on TOP.' (i.e., He WON the game.)

2.1.2 Akan

In Akan, the construction for both SF and NSF is characterized by two features. It obligatorily makes use of the FM *nà* and displays a so-called "link tone" (Bearth 2002; cf. Schachter 1973 as well) at the verb in the out-of-focus part.

In SF, an expletive subject pronoun for 3rd person referents (*ὲ-*) is characteristically used, although it might be replaced by the normal pronominal form.[3] In NSF on the other hand, there is no general need for an object pronoun that is coreferent with the constituent in focus, but rather the selection of the object pronoun underlies semantic criteria. Animates require a pronoun (cf. 4a), especially when human, inanimates do not allow it. In the absence of a pronominal object, the "past" transitive verb in sentence-final position carries suffix *-yɛ* (cf. 4b).

---

[3]   Bearth et al. (2002) describe the change of the subject pronouns as restricted to human referents. This distribution is not supported by our data.

SF

(3)   ɛ̀-yɛ̀   àbrɛ̀wá   nó   nà ɛ̀-díì   àdùá nò.
       3sg-COP old.woman DEF NA 3sg-eat beans DEF
       'It is the OLD WOMAN who ate the beans.'

       cf. the canonical sentence

       àbrɛ̀wá   nó   dìì   àdùá nò.
       old.woman DEF   eat   beans DEF
       'The old woman ate the BEANS.' ~ '... woman ATE THE BEANS.'

NSF

(4) a.   nè krámán nà pàpá nò   súà   nò.
          3sg dog   NA man DEF carry O:3sg
          'The man carried his DOG.'

     b.   nè bágè nà ɔ̀-súà-yɛ́.
          3sg bag   NA 3sg-carry-YE
          'He carried his BAG.'

2.1.3 Lelemi

In Lelemi, the difference between subject and non-subject focus constructions lies above all in the verbal morphology. Lelemi has two sets of TMA-markers: one used in simple tenses, the other in so-called "relative" tenses (Allan 1973). Not every simple tense has a counterpart in the relative tenses.

The "relative tenses" ("relative past, relative present, relative future and relative present for verbs of state") show up in SF constructions. Unlike the simple tenses, they don't have noun class concord for the subject, and they display only one invariant form. The simple tenses, on the other hand, are used in NSF. They are formed by subject noun class concord markers, often assimilated with the following segmentally and/or tonally coded TMA morphemes, and the verb (with grammatical tone for each tense).

The morphological coding device for NSF consists of morpheme *nà* postponed to the focused constituent (cf. the homophone FM in Akan). Some of the informants treated it as obligatory, others claimed that the syntactic marking, i.e. the sentence-initial position, makes already clear that this constituent is in focus.

SF   cf. verb form in
                                                                      simple tense:

(5) a.

 ɔ̀nàabì ùmwì pɛ́ mɔ̀-dì   kùtú.   ɔ̄ɔ̄-dì

boy   one   only REL.PRS(dyn)-eat orange
'Only ONE boy is eating an orange.'

   b. ɔ̀nàabì n̄-t̄ī   ùlòkúbì.   ū-t̄ī
boy   REL.PRS(stat)-carry girl
'A BOY was carrying a girl.'

   c. lɔ́ɔ́lì ínyɔ́ ná-sà.
                                                          lɛ́-sà

lorry two   REL.PAST-meet
'TWO LORRIES COLLIDED.'[4]

NSF

(6)   àkábí áwɔ̄dí (nà) ùlōkū ɔ́mɔ̀ ɔ́-dì.
       beans raw   (NA) woman DEF   3sg.PAST-eat
       'The woman ate RAW BEANS.'

---

[4] In Lelemi, SF and sentence focus are coded in the same way – a feature which is in fact characteristic for Gur languages.

2.1.4 Buli

Buli has a preposed affirmative FM *ká* (negative suppletive *dāā*) in SF as well as in NSF. The affirmative morpheme seems to be optional in both constructions. In SF, the focus constituent is always followed by the conjunction *lē*, while in NSF we typically find the conjunction *tè*.[5] In NSF, the occurrence of *lē* is less common, but not totally excluded.

Verb tone deviates from that in simple sentences in both focus constructions, although not in the same way. In fact, Buli has three paradigms distinguished by grammatical tone on dynamic perfective verbs in clause-final position: a canonical paradigm A in simple sentences, a paradigm B after conjunction *lē* (SF) and a paradigm C that shows up in clauses with the conjunction *tè* (NSF).[6] In both constructions, pronouns which are coreferent with the focused constituent do not occur in the out-of-focus part.

SF

(7)   (ká)   wá   lē   chèŋ.   Paradigm B, not: *tè*
       (FM)   3sg:DISJ   CNJ   go
       'HE went.' ~ 'It is HE who went.' (not you)

---

[5] Both conjunctions are sometimes provided with an initial vowel (*àlē, àtè*). This vowel occurs with other clause-inital conjunctions as well as with clause-initial serialized verbs and is always correlated with a prosodic break before the clause.

[6] Paradigm B is characterized by an "instabil rising tone" (Schwarz 2004: 38) and paradigm C by an invariable low tone. Both paradigms are constituted by the absence of subject congruent grammatical tone operating in simple clauses (Paradigm A). The neutralization in paradigm C versus A shows up only with discourse participants (1st and 2nd person), while it is not evident with 3rd persons, as in ex. 7-8.

cf.   the canonical sentence

wà   chèŋ-kà   sándēm.   Paradigm A
3sg   go-FM[7]   Sandema
'He went to SANDEMA..' ~ 'He WENT TO SANDEMA.'

NSF

(8) a. (ká) sándēm tè   wà chèŋ.   Paradigm C, rare:
         lē
         (FM) Sandema CNJ 3sg go
         'It is SANDEMA where he went.' (not Navrongo)

     b. john tè   mí   fɔb.
         John CNJ 1sg:DISJ slap
         'I slapped JOHN.'

2.1.5 Dagbani

Dagbani provides a strong structural asymmetry between SF and NSF. The ca-
nonical SF construction[8] contains a postponed syllabic nasal called "emphatic"
by Olawsky (1999). In NSF on the other hand, the so-called FM *kà* (Olawsky
1999: 63) has to be put at the beginning of the out-of-focus part.

    Similar to Buli, both focus construction types are distinguished from each
other by grammatical verb tone and deviate from the verb paradigm in simple
sentences. A pronominal form for a focused subject is not allowed and in
general, neither focused non-subjectival verb arguments nor other focused
constituents are pronominally represented in the out-of-focus part.

---

[7] Please note that in the canonical indicative sentence in which either the complement or the
     whole VP might be focal, FM *ká* is enclitizised to the verb (the morpheme's surface tone
     may change). In *ex-situ* focus constructions any postverbal position of this morpheme is
     completely excluded.
[8] There is another SF construction formed with post-subjectival *léé* which is however almost
     restricted to questions and not considered here.

SF

(9)   páɣá   má!á   m̀   bɔ́l-ɔ̀.   Paradigm B
       woman DEF   N   call-O:3sg
       'The WOMAN called him.'

       cf. the canonical sentence

       ɔ̀   bɔ̀l-lá   george.   Paradigm A
       3sg call-FM[9] George
       'She called GEORGE.' ~ 'She CALLED GEORGE.'

NSF

(10)   george   kó-!ó   bɔ̀lì.   Paradigm C
       George   KA-3sg call
       'She called GEORGE.'

2.2   Relative Structures

Concerning the often stated "relative" forms in *ex-situ* focus constructions, we found that relative structures, i.e. a head and a (restricting) relative clause, are not present in all of our sample languages. And if they exist in *ex-situ* focalization, they are not necessarily identical with the language specific prototypical RCs with restrictive reading, as demonstrated in the following.

2.2.1   Ewe

Ewe disposes of a general construction to express restrictive relative clauses. The overall features of this construction are: (i) the relative clause is introduced by a demonstrativum *sì* (standard-Ewe) or *yíkè* 'this' (dialectal variant for Inland Ewe) which takes over the function of a relative pronoun and (ii) it ends

---

[9]   In case of complement or VP focus, suffixed FM *-lá* occurs in Dagbani indicative sentences. Like Buli FM *ká* in postverbal position, this suffix is excluded in *ex-situ* focus constructions.

generally with the determiner *lá*. If the relative clause follows the main clause, the determiner is sometimes omitted.

There is a difference among relative constructions depending on the syntactical function of the antecedent: if it is the subject of the relative clause, no pronominal form is used, if it is a non-subject then we find the already mentioned special pronouns which only show up in 2nd and 3rd person singular.

Subject RC

(11) éyé tsīitsītɔ̄   sì kɔ́ èlā̀ lā vê̄   lá, trɔ́ dzó ...
      CNJ elder.one   DEM take fish DET come.to DET, change return
      'And the elder one, who brought the fish, returned back ....'

Non-Subject RC

(12) a.   nú sì wò-   gblɔ̀ ná mì lá   mì wɔ̀ è.
          thing DEM 3sg.DEP say for 2pl DET 2pl do O:3sg
          'What he tells you, do it!' (Duthie 1996: 45)[10]

     b.   ńtsùvǐ ɖé tɔ́   ɖé
          boy   IND stand reach
          fí yíkè wó- nɔ̀ bɔ̀l ɸōḿ   lè   ɸé gôl mè.
          place DEM 3pl HAB ball beat. PROG be.LOC place goal in
          'A boy is standing in a goal.' (lit.: 'A boy is standing at the place where they normally shoot the ball into.')

There are only minor similarities between focus and relative constructions. These concern the use of the pronominal forms: as in SF, in Subject RC no pronoun is used, whereas in NSF as well as in Non-Subject RC the dependent subject pronouns are found. The two features of RC mentioned above are not

---

[10] The interlinearization was done by the authors.

present in focus constructions, and, on the other hand, morpheme (*y)é* of focus constructions is absent in RCs.

2.2.2 Akan

Akan relative constructions are characterized by the use of (i) the so-called "relative pronoun" *â* and (ii) a clause-final determiner *nó*. Furthermore, the verb in the relative clause changes its tone pattern in adopting a H tone (cf. Schachter 1973, the so-called "link tone" in Bearth 2002).

Subject RC

(13) <u>àbrántìè n(o)-áà ɔ-bɔ́ɔ̀ wó nó</u>, yɛ̀ m-àdámfùɔ̀.
       boy   DEF-REL 3sg-hit O:2sg DET   COP 1sg-friend
       'The boy who hit you is my friend.'

Non-Subject RC

(14) <u>àbrántìè n(o)-áà wò-bɔ́ɔ̀ nò   nò</u>, yɛ̀ m-àdámfùɔ̀.
       boy   DEF-REL 2sg-hit O:3sg DET   COP 1sg-friend
       'The boy whom you hit is my friend.'

Hence, in relativization two additional features show up compared to focus constructions. On the other hand, post-focal morpheme *nà* is missing in RC. What both constructions share in comparison to simple sentences is only the "link tone".

2.2.3 Lelemi

Lelemi displays a subject- and non-subject asymmetry in relative constructions that resembles the dichotomy in its *ex-situ* focus constructions (cf. 2.1.3) and that is based on the syntactic function of the relativized element within the RC. Common component in both types of relative constructions is the determination of the relativized element by an identifier pronominal form that consists of a noun class concord for the preceding relativized noun and morpheme -*ní*. If the subject is relativized, a "relative" TMA form, i.e. the TMA form without subject

prefix at the verb is used (cf. the simple verb form in the initial clause of 15 with the relativized below). If a non-subject constituent is relativized, the simple TMA verb form, i.e. including subject prefix, occurs.

Subject RC

(15)   èbùò   ū-ɲè.
         animal 3sg.PRS-stand
         'There is an animal there,

         ńzù   èbùò   ɔ́-ní   n̄-ɲè   vɔ̄   ɔ́-ɗì   ɔ̀náànjùé.
         but   animal 3sg-NI   REL.PRS(stat)-stand there 3sg-COP cow
         but the kind of animal that is over there is a cow.'

Non-Subject RC

(16)   ɔ̀klàmá   ɔ́ŋvɔ̀ ɔ̀ɔ̄-dù   trouzis ɔ́-ní   ɔ̄nāābì ɔ́ŋvɔ̀ ɔ́-chà
         dog   DEM 3sg.CONT-bite trousers 3sg-NI boy   DEM 3sg.PAST-wear
         'The dog bites the trousers which the boy wears.'

Apart from the use of the "relative" versus the simple TMA-forms, *ex-situ* focus constructions and RCs with restrictive meaning are not the same. It is first of all the "identifier pronoun" (noun class pronoun *–ni*) which distinguishes the restrictive RC from focalization. A further element that is typical for NSF contrary to relativization is morpheme *nà* at the beginning of the out-of-focus part.

2.2.4 Buli

Buli disposes of two structural types of RC[11] which share the following features: (i) the relativized head is provided by an indefinite noun class pronoun[12], like

---

[11] Cf. Hiraiwa 2003 for a detailed description of RCs in Buli.
[12] The indefinite pronouns refer to specific indefinite entities and are therefore translated as 'certain, some' in other contexts.

wāā(i), which can either represent the head on its own or forms a nominal compound with an initial nominal constituent, and (ii) determiner *lá* is commonly added to the end of the relative clause. This determiner is sometimes omitted in sentence-final RCs. The head internal RC type shares features with SF constructions, since it consists of preverbal conjunction *lē*. It is always used when the relativized element has subject function in the relative clause and it can be found with relativized verb objects, too[13]. The other RC is of the head-external type and not compatible with a subject as head. Structurally, this RC resembles NSF, since it contains conjunction *tè* (ex. 18b). The grammatical tone of perfective verbs deviates in both RC types from paradigm A with tonal subject agreement.[14]

Subject RC

(17)   <u>núrú-wāā   lē chèŋ lā</u> ká mí   dɔ̄ā.
       person-IND:CL CNJ go   DET COP 1sg:DISJ friend
       'The person who has left is my friend.'

Non-Subject RC

(18) a.   <u>fì lē yàlì nùrú-wāā   lá</u> ká mí   dɔ̄ā.
          2sg CNJ marry person-IND:CL DET COP 1sg:DISJ friend
          'The person you married is my friend.'

     b.   <u>núrú-wāā   tè fì yàlì lā</u> ká mí   dɔ̄ā.
          person-IND:CL CNJ 2sg marry DET COP 1sg:DISJ friend
          'The person you married is my friend.'

---

[13]   The head-internal RC therefore has either a postnominal (relativized subject, cf. ex. 17) or a circumnominal (relativized object, cf. ex. 18a) structure.

[14]   In RC with clause-final determiner *lá* there is however no "instabil rising tone", since this paradigm B pattern is restricted to perfective verbs in clause-final position (cf. clause-final *chèŋ̄* in ex. 7 versus ex. 17).

Despite the similarities in morphology (conjunctions *lē* and *tè*) and tonal neutralization (no paradigm A for dynamic verbs in perfective), both relative clause types with restrictive reading contain components which distinguish them from the *ex-situ* focus constructions: the indefinite class pronoun and the RC-final determiner.

2.2.5 Dagbani

Similar to Buli, Dagbani has two RC types at its disposal which share the following features: (i) the head is represented by an indefinite noun class pronoun (like *so* in examples (19) and (20))[15] that forms a compound with the nominal root or is used alone, and (ii) determiner *máá* (sometimes *lá*) is added to the end of the RC.

One of the RC types is restricted to cases in which the antecedent has subject function within the RC (ex. 19). Apart from the two features mentioned above, it is formed with the help of a disjunctive pronoun in the subject slot which follows the head and fulfills here the function of a relative pronoun.[16] The other RC type occurs only with non-subjects as relativized heads and makes use of post-subjectival particle *n(i)* marking also some other subordinated clauses. In this head-internal RC type the head is either retained in its postverbal slot (cf. the circumnominal ex. 20a) or it is moved to the initial position of the relative clause (cf. the postnominal ex.20b).

Subject RC

(19) <u>dò-só</u>   ŋún   !cháŋ má!á ɲέ-!lá   ń   zò.
      man-IND:CL 3sg.DISJ go   DET   COP-FM 1sg friend
      'The man who has left is my friend.'

---

[15] The indefinite pronoun consists of a stem *s[front vowel]* (the vowel occurs only in case of CV suffixes) which is provided with a class suffix. It corresponds functionally to the indefinite pronoun in Buli (cf. footnote 12), although the latter lacks the *SV* stem.

[16] According to Wilson (1963: 139), the indefinite pronoun is optional in subjectival relative clauses.

Non-Subject RC

(20) a. <u>á-n(í) !ŋmé dò-só   má!á</u> ɲɛ́-!lá   ń   zò.
         2sg-CNJ hit   man-IND:CL DET   COP-FM 1sg friend
         'The man whom you have hit is my friend.'

    b. <u>dò-só   á-n(í) !ŋmé má!á</u> ɲɛ́-!lá   ń   zò.
        man-IND:CL 2sg-CNJ hit   DET   COP-FM 1sg friend
        'The man whom you have hit is my friend.'

The comparison with the focus constructions shows no direct correspondence, but it is possible that the subordinating particle *n(i)* and the "emphatic" syllabic nasal *N* that obligatorily occurs in SF are related to each other.

2.3 Narrative Structures

Our third observation concerns the fact that there is a constant formal parallelism between *ex-situ* non-subject-focus constructions (NSF) and narrative clauses (NC), and in some of our sample languages the narrative structure is extending to SF, too. We use the term NC for clauses that encode the succession of events in realis mood and that serve to continuitively develop the main story line. Labov regards this function as follows: "Each clause … describes an event that is understood to shift reference time, i.e. it follows the event immediately preceding it, and precedes the event immediately following it." (1972, cited in Schiffrin 1994: 284).

The formal parallels show up in several ways in the selected languages and are realized by the FM, TMA morphology including tone, and special pronominal forms.

2.3.1 Ewe

Ewe shows certain structural similarities between NSF and NC, although there is no total correspondence. These are best demonstrated by the use of special subject pronouns[17] in both constructions, as can be seen in ex. (21) and (22).

Additionally, there is a similarity of the FM with a clause coordinating conjunction *éyē* 'and (then)' which is found in narrative contexts.

NSF

(21) èdzì(-é)   wò-   ɖù.   (= ex. 2)
      top(-YE)   3sg.DEP   eat
      'He was on TOP.' (i.e., He WON the game.)

NC

(22) And the third one .. found the way to the market …

      yá   wò-   vá   kpɔ̄   tòmátós.
      CNJ[18] 3sg.DEP come   see   tomatoes
      '… and he found tomatoes.'

2.3.2 Akan

Akan has a clausal sequential conjunction *nà* with the meaning 'and (then)' (Bearth 2002) which is identical with the FM. The verbal morphology including the "link tone" also seems to be the same in both clauses, although this is still a matter of further research.[19]

---

[17] This has already been noted by Duthie (1996: 53) and Ameka (2004: 17).
[18] The form *yá* is one dialectal variant of the conjunction *éyē* in Inland Ewe.
[19] Bearth (2002) postulates the existence of a "link tone" on the verb as well as the existence of the so-called "dependent" morpheme *yɛ* only in *ex-situ* focus constructions, while our own data exhibit them in other contexts, too, including sequential events with clause-initial conjunction *nà*.

NSF

(23) àdùá   nà   ɔ̀-díî-yɛ́.
beans NA   3sg-eat-YE
'He ate BEANS.'

NC

(24) máàmé nòáà àdùá,   nà   n-àdámfùò díî-yɛ́.
Maame cook   beans CNJ 3sg-friend eat-YE
'Maame cooked beans and her friend ate them.'

2.3.3 Lelemi

In Lelemi, NSF and NC clauses show identical features: In both, the simple tenses are used. Furthermore, the FM is homophone with the narrative conjunction 'and' which coordinates two clauses and we suppose that it is the same morpheme. It is segmentally identical with the "relative past" tense morphem, too (cf. ex. 5c).

NSF

(25) àkábí áwɔ̄dí (nà) ùlōkū ɔ́mɔ̀ ɔ́-dì.   (= ex. 6)
beans raw   (NA) woman DEF 3sg.PAST-eat
'The woman ate RAW BEANS.'

NC

(26) 'The youngest child went …'

nà   ú-tī   ùlū gɛ̃̀ɛ̃̀ ɔ́mɔ̀.
CNJ 3sg.PAST-take road right DEF
'and he took the right road.'

2.3.4 Buli

Buli, too, displays a striking parallel between NSF and NC. First, the clause initial element *tè* of the out-of-focus part of NSF and the clausal conjunction *tè*

'and' in NC[20] are completely identical. Second, the identity pattern stretches onto the verb: it bears the same grammatical tone (paradigm C) after *tè* in both functions and it differs thus from (i) the canonical paradigm A, and (ii) the paradigm B that is found after marker *lē* (i.e. mainly with SF, cf. 2.1.4.). This is true only for the group of dynamic verbs in perfective aspect as shown in the following examples. Dynamic verbs in the imperfective as well as stative verbs do not participate in paradigm C, but return to paradigm A where they are most often – although not throughout – marked as irrealis, i.e. the non-perfective verb forms tend to occur in subjunctive in focus constructions.

NSF

(27) (ká) sándēm   tè   wà   chèŋ.   (= ex. 8a)
      (FM) Sandema   CNJ   3sg   go
      'It is SANDEMA where he went.' (not Navrongo)

NC

(28) and his mother was happy with him

      tè   bà   dìg   jèntàŋā ...
      CNJ   3pl   cook   soup.DEF
      'and they cooked the soup ...'

2.3.5 Dagbani

Finally, Dagbani also affirms the parallel pattern between NSF and NC convincingly. Morpheme *kà* which is following the focused constituent in NSF constructions has a clausal conjunction counterpart *kà* 'and' in narrative contexts. Furthermore, the grammatical tone of dynamic verbs in such clauses differs in the same way from the subject congruent verb tone in simple clauses,

---

[20] At the beginning of subjunctive clauses (irrealis), the conjunction is also consecutively used.

irrespective whether we deal with a real narrative context or with a focus construction.

The coding of the second clause in ex. (30b) demonstrates Olawskys (1999: 44) observation that, if the subject of the clause introduced by *kà* is coreferent with the subject of the preceding clause, it has to be elided.

NSF

(29) yílɔ́   má!á   nì   kó-!ó   dì.
      house DEF   in   KA-3sg   eat
      'In the house she ate.'

NC

(30) a.   and the mother sent the youngest child

          kà   bíí   má!á   chàŋ ...
          CNJ   child   DEF   go
          'and the child went ...'

     b.   páɣá   má!á   dáá-!lá   peter   ká   !ŋmé-ò.
          woman DEF   push-FM   Peter   CNJ   hit-O:3sg
          'The woman pushed and hit Peter.'

                              not: ... ka *o ŋme-o.

2.4   Diversity and Distribution of Forms

Summarizing our observations concerning the structural features of affirmative *ex-situ* focus constructions, relative clauses and narrative clauses, we have to state that the *ex-situ* focus constructions minus the focus constituent itself (F) resemble relative clauses only to a certain extent while the structural features of narrative clauses are matched much closer. Table 1 gives an overview on the differences and parallels as they have emerged in our investigation:

Table 1

	Kwa Languages			Gur Languages	
	Ewe	**Akan**	**Lelemi**	**Buli**	**Dagbani**
SF	F *(y)é*	*(ɛ̀yɛ̀)* F *nà* expl. *ɛ̀* link tone	F rel. tense	*(ká)* F *lē* tone B	F *N* tone B
RC₁	*sì~(yí)kɛ̀* DET *lá*	*â* link tone DET *nó*	CL-*ni* rel. tense	IND:CL *lē* DET *lá*	IND:CL *n(i)* DET *la~máá*
RC₂	*sì~(yí)kɛ̀* DEP pron. DET *lá*	--	CL-*ni* simple tense	IND:CL *tè* DET *lá*	IND:CL DISJ pron. DET *la~máá*
NSF	F *(y)é* DEP pron.	*(ɛ̀yɛ̀)* F *nà* link tone	F *nà* simple tense	*(ká)* F *tè* tone C	F *kà* tone C
NC	... *éyē* DEP pron.	... *nà* link tone	... *nà* simple tense	... *tè* tone C	... *kà* tone C

With respect to RCs, we face considerable heterogeneity among our sample since the two Gur languages provide especially strongly divergent RC types. This is due to the additional head-internal relative clause type in these two languages which is represented by RC₁ in table 1 and which seems structurally related with the SF constructions in both languages (cf. preverbal morpheme and grammatical tone paradigm). The head-external RC type among these languages is represented in the table by RC₂, but only in Buli it also displays evident formal parallels with a focus construction, namely with NSF. Considering the relation between RCs and the non-focal part of *ex-situ* focus constructions in the Kwa languages, we again note only partial correspondences: in Ewe, dependent pronouns occur in RCs when a non-subject is antecedent as well as in NSF; in Akan, tonal changes ("link tone") pertain in RC as well as in SF and NSF; and in Lelemi, the selection of the tense form in RC and focus constructions is due to the syntactic function of the preceding relativized respectively focused element as subject or non-subject. Absent in focus constructions are however those morphological means which are characteristic for almost all restrictive RCs

throughout our language sample, i.e. the respective (relative / identifying / indefinite) pronominal forms accompanying the head and, with the exception of Lelemi, the RC-final determiners.

By integrating NCs into our considerations, it becomes evident that part of the so-called "relative" features in focus constructions are not just exclusively "relative", if at all, but should rather be analysed as structural reflections between coordinated "narrative" clauses. Contrary to the complex picture with regard to the RC pattern in focus constructions, all five Ghanaian Gur and Kwa languages considered here display in fact a very close correspondence between (N)SF and NC. With the exception of Ewe and some need for verification in Akan, we can even postulate a complete structural identy for both. We therefore conclude that the parallelism between the out-of-focus part of morpho-syntactically marked (N)SF and narrative clauses (NC) is no coincidence, but is due to a systematic "narrative" basis of the respective focus constructions.

3   Narrative Hypothesis

From the structural distribution above it is evident that the parallelism between (N)SF and NC is a systematic pattern. We propose that in fact a narrative clause constitutes the non-focal part of such *ex-situ* focus constructions and that its grammaticalization can conceal their biclausal structure. We therefore don't follow the movement hypothesis as for example suggested by Aboh (2004) and Green (1997) for Gbe and Hausa respectively. Like in the cleft hypothesis assumed for various languages (Givón 1990/2001, Schachter 1973, Heine/Reh 1984, Lambrecht 2001), our proposal considers the constructions as biclausal and adds a diachronic perspective to their synchronic analysis. Unlike in the prototypical cleft hypothesis however, we here assume a NC rather than a RC as source for the potential or the already undergone grammaticalization processes. In the following, we will argue for the validity of the narrative hypothesis for each language separately.

3.1   Ewe

Ewe focus constructions can be regarded as derived from two coordinated clauses, where the second clause is provided by an original narrative clause while the first clause is commonly represented by an NP alone. A copula form is only needed in the latter, if the focused constituent is negated, a fact which holds for all the languages considered here.

The synchronic FM *yé* shows structural similarities to the conjunction *éyē* (which is underlying /*éyé*/ [21]) 'and, and then', although it is not identical. The meaning of the conjunction's prefix *é-* hasn't been explained convincingly up to now. We assume that the FM has developed out of the conjunction. Following this grammaticalization path, one has to claim a divergent development of the synchronic conjunction and FM. The latter is in normal speech usually eroded to vowel *-é* and suffixed to the preceding NP, that is, it has become part of the initial clause.

Synchronically, Ewe displays a homophone morpheme *yé* occurring in nominal predication like 'It is a pen.' – *pên yé*. Here, its function is comparable to a copula verb.

A further structural feature supporting the narrative hypothesis in Ewe is the use of the special subject pronouns in NSF and NC. Westermann (1930: 61) mentions that they are used "in the continuation of a sentence, or closely to connect one sentence with a preceding one." It is only in subject focus constructions, where they are not required and even ungrammatical. If we assume a narrative construction underlying both, SF and NSF, then we have to state that in SF the structure of the original source is extensively eroded due to a phenomenon we will call here the "double-subject" constraint, i.e. the focused constituent cannot be followed by a coreferential pronominal subject in SF.

---

[21] Cf. Clements (1977: 172) for the tone rules changing the two phonological high tones of the conjunction.

3.2 Akan

The first hint for the validity of the narrative hypothesis in Akan is the identity of FM and narrative conjunction. Our informants treated *nà* in FC still as conjunction so that, if there is a certain degree of grammaticalization at all, as suggested by its description as FM by some authors, this could be only by a functional split in the very inceptive stage.

The first clause of the biclausal focus construction is often only represented by an NP. Alternatively, the initial clause starts with *èyè*, i.e. an expletive pronoun plus copula verb 'to be' (cf. ex. 3).

Different from Ewe, in Akan, the biclausal status of the subject focus construction is still well maintained, since the out-of-focus clause obligatorily requires a subject pronoun, which might be an expletive one (cf. 2.1.2.). The "double subject" constraint is thus not operating in this language.

3.3 Lelemi

In Lelemi, the narrative clause as part of an *ex-situ* focus construction is evident on first sight only for NSF. As shown above, the non-focal clause of NSF and the narrative clause are formally totally identical, i.e. any probably assumed grammaticalization of the narrative clause is restricted to the functional level and has no effects on the structural level. Accordingly, morpheme *nà* is in both functions considered as conjunction by us.

In SF, the conjunction is missing and "relative" tense forms are used instead of the simple tense forms. Nevertheless, we can assume such grammaticalization source in one of the relative tenses/aspects. The TMA morpheme for the "relative past" tense is high toned *ná*. We analyse it as a conglomeration of the conjunction *nà* (with inherent Low tone) plus a High tone which is born by the subject pronoun in simple past. The slot for the pronoun is not filled due to the "double subject" constraint. The high tone it bears in simple past is however retained with the former conjunction.

3.4 Buli

In Buli, the narrative hypothesis is valid for the prototypical NSF construction which is formed with conjunction *tè* and tone paradigm C. Since these features are shared by sequential clauses in narrative contexts, too, a narrative clause can be regarded as representing the non-focal part in NSF. The lack of tone paradigm C with stative verbs as well as the frequent modal change of dynamic verbs in the imperfective (cf. 2.3.4.) supports the proposed narrative status of the respective clause.

The SF construction on the other hand requires conjunction *lē* which cannot be related to the narrative conjunction as such, but is rather segmentally identical with the NP coordinating conjunction *lè* 'and, with'. This structural similarity among the two "*le*-type" junctors and the very strict "double subject" constraint in SF might be an indication for a semantico-syntactically closer conceptualization of this construction as one single information structural unit compared to the evidently biclausal NSF organization with *tè*.

3.5 Dagbani

Like in Buli, the grammaticalization of the NC clause is restricted to NSF in Dagbani. The so-called "focus marker" *kà* is in fact just a conjunction at the beginning of a NC clause which requires verb tone paradigm C (for dynamic verbs). In NSF, a subject constituent must always follow the morpheme *kà* while coreferent subjects in two sequential clauses via clause junction *kà* are ungrammatical (cf. 2.3.5.). Since there are no constitutive structural differences between the non-focal part in NSF construction and the basic NC clause, *kà* is in both contexts still analyzed as conjunction by us, though it has some potential for grammaticalization into FM.

Interesting is the parallel to Buli found in Dagbani, insofar as here the "emphatic" marker *N* in SF resembles the NP coordinating conjunction *nì~ǹ* 'with, and'. It seems that Dagbani has a similar tendency towards a closer and

more intraclausal organization of SF compared to NSF and hence does not make use of the narrative pattern with biclausal coordination in SF.

4   Comparative Summary

As we have shown, there are striking similarities on the morphosyntactic level between the non-focal part of focus constructions and NCs, although the relevant structural parameters diverge even in our small language sample due to typological subtraits. Hence, in some of the languages – namely Akan, Buli, and Dagbani – grammatical verb tone must be taken into account in order to identify the non-focal part of focus constructions as NC. All of the languages make use of special morphological means. Apart from clausal conjunctions this also concerns suppletive pronouns. Not surprisingly it is Ewe, a language known for its pronominal specialization including logophoric forms, that provides the NC and the focus construction based on it with more than just one "dependent" pronominal form.

A typologically interesting picture in our small language sample is displayed by the distribution of the narrative structures as such in *ex-situ* focus constructions, as shown in table 2.

Table 2

On the one hand, in the Kwa languages studied, including Lelemi, the narrative pattern is more or less overtly extending into SF constructions. In the two Gur languages studied here on the other hand, SF constructions do not participate in the narrative pattern. Schwarz (in preparation) shows that in languages of this group SF rather tend to be represented by a syntactically more hierarchical (head

internal) relative construction and that the distribution of the two *ex-situ* focus constructions can be accounted for on discourse organizational grounds taking the notion of topic into account.

Having concentrated here on a deeper insight into the narrative structures that have emerged in (N)SF throughout our sample, we claim that a clausal conjunction as used to coordinate sequentials in narration does also function as device to link together focus constituent and non-focal part in a non-hierarchically way. Such focus constructions are consequently to be considered as basically biclausal, even if the clausal status of the initial clause with the focused constituent is not reflected throughout. In some of the languages, the inceptive stages of grammaticalization processes of the clause-initial conjunction into FM can be perceived, a grammaticalization chain that may even stretch further into a copula-like predicative morpheme as noticed by Stassen (1997: 85). The actual stage of such grammaticalization chain in our sample languages is shown in table 3:

Table 3

	CNJ →	FM →	COP
Ewe	*éyē*	*yé, -é*	*yé*
Akan	*nà*	*(nà)*	--
Lelemi	*nà*	*(nà)*	--
Buli	*tè*	--	--
Dagbani	*kà*	*(kà)*	--

As can be seen in the table, Akan and Lelemi exhibit homophone morphemes which could be a result of borrowing from Akan to Lelemi since loans from Akan are common in the Togo mountain area.

Three of the languages, namely Akan, Lelemi and Dagbani display the same pattern insofar as they have a conjunction which has been interpreted by some authors (Boadi 1974, Ameka 1992, Olawsky 1999) as a right-adjacent FM. According to us, the respective morphemes do have the potential for such a

functional split, but that stage has probably not yet been reached, since we could not notice any relevant categorial or structural changes of the conjunction towards a FM.

As for the Buli conjunction, there are no indications at all that it might take the grammaticalization path into a focus or predicative marker in the near future. Responsible for that is first its restriction to NSF, a fact that the Buli conjunction shares with the respective Dagbani conjunction. Second, the Buli clause conjunction is in affirmative focal contexts relatively often counter-balanced by the predicative marker respectively FM *ká* left to the focus constituent, while such an affirmative counterpart is missing in Dagbani. If the focused constituent is negated, all five languages make however use of negative copula forms. We conclude that the rarer the copula forms in affirmation are, the higher are the chances for reanalysis of the clausal conjunction as FM.

Contrary to the rather inceptive stage of grammaticalization if existent at all in most of the languages, there seems to have been a longer development in Ewe. Here, the original conjunction already shows signs of erosion when functioning as FM and it is even often suffixed to the constituent in focus.

As noted in 3.3., in Lelemi the conjunction *nà* has taken a special direction in grammaticalization. Together with the high tone born by the subject prefix in other syntagmata, it has become a "relative past" tense marker in SF. Such development from a conjunction denoting the accomplishment of actions to a past marker was also shown by Hopper (1979) for Malay, an Austronesian language.

Table 4

	CNJ $\rightarrow$	**"Relative Past"**
Lelemi	*nà*	*ná* ($\leftarrow$ *nà* + ´)

Our aim here was to defend the existence of a steady narrative pattern in *ex-situ* focus constructions and to outline the diachronic implications of the narrative hypothesis as an alternative to cleft and movement approaches. It has been

shown that not only the linguistic expression of the in-focus part, but also that of the out-of-focus part is important for an adequate analysis of *ex-situ* focus constructions and that the functional load verb morphology including tone has in African languages in this respect should not be underestimated.

Reference

Aboh, Enoch Oladé. 2004. *The morphosyntax of complement-head sequences. Clause structure and word order patterns in Kwa:* Oxford studies in comparative syntax. Oxford, New York: Oxford University Press.

Allan, Edward Jay. 1973. *A grammar of Buem: The Lelemi language,* School of Oriental and African Studies, University of London: PhD Thesis.

Ameka, Felix. 1992. Focus constructions in Ewe und Akan: A comparative perspective. In *Proceedings of the Kwa comparative syntax workshop* M.I.T., eds. Chris Collins and Victor Manfredi, 1-15. Cambridge.

--   2004. Grammar and cultural practices: The grammaticalization of triadic communication in West African languages. *Journal of West African Languages* 30:5-28.

Bearth, Thomas. 1993. Satztyp und Situation in einigen Sprachen Westafrikas. In *Beiträge zur afrikanischen Sprach- und Literaturwissenschaft,* ed. W.J.G. Möhlig, 91-104. Köln.

--   2002. Fokus im Akan - Outline. Handout of a presentation at the Linguistic Colloquium of the Seminar of African Studies at the Humboldt-University, Berlin, 28.05.2002.

Bearth, Thomas, et al. 2002. Ali Akan. Version 56. CD-ROM.

Blench, Roger. 2001. Comparative Central Togo: what have we learnt since Heine? Paper presented at *32nd Annual Conference on African Linguistics,* Cambridge.

Boadi, Lawrence A. 1974. Focus-marking in Akan. *Linguistics* 140:5-57.

Clements, George N. 1977. Four tones from three: The extra-high tone in Anlo Ewe. In *Language and linguistic problems in Africa. Proceedings of the VII. conference of African linguistics,* eds. Paul F. A. Kotey and Haig Der-Houssikian, 168-181. Columbia: Hornbeam Press.

Duthie, Alan S. 1996. *Introducing Ewe linguistic patterns.* Accra: Ghana University Press.

Fiedler, Ines, and Schwarz, Anne. (to appear). Focus or Narrative Construction? In *Focus Strategies: Evidence from African Languages*, eds. Enoch Aboh, Katharina Hartmann and Malte Zimmermann. Berlin: de Gruyter.

Frajzyngier, Zygmunt. 2004. Tense and aspect as coding means for information structure: a potential areal feature. *Journal of West African Languages* 30:53-67.

Givón, Talmy. 1990. *Syntax. A functional-typological introduction.* Vol. 2. Amsterdam: John Benjamins.

Green, Melanie. 1997. Focus and copula constructions in Hausa, School of Oriental and African Studies, University of London: Ph.D.Diss.

Heine, Bernd, and Reh, Mechthild. 1984. *Grammaticalization and reanalysis in African languages.* Hamburg: Helmut Buske.

Hiraiwa, Ken. 2003. Relativization in Buli. In *Studies in Buli Grammar. Working Papers on Endangered and Less Familiar Languages*, ed. Michael Kenstowicz, 45-84: MIT Working Papers in Linguistics.

Hopper, Paul J. 1979. Some observations on the typology of focus and aspect in narrative language. *Studies in Language* 3:37-64.

Hopper, Paul J., and Traugott, Elizabeth Closs. 1993. *Grammaticalization.* Cambridge: Cambridge Univ. Press.

Labov, Wilhelm. 1972. The transformation of experience in narrative syntax. In *Language in the Inner City,* 354-96. Philadelphia: University of Philadelphia Press.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39:463-516.

Olawsky, Knut J. 1999. *Aspects of Dagbani grammar. With special emphasis on phonology and morphology*: LINCOM studies in African linguistics; 41. München, Newcastle: Lincom.

Rongier, Jacques. 1997. Langues autonomes du Togo, entre Gur et Kwa?, Ms.

Schachter, Paul. 1973. Focus and relativization. *Language* 49.

Schiffrin, Deborah. 1994. *Approaches to discourse*: Blackwell Textbooks in linguistics. Oxford UK Cambridge USA: Blackwell.

Schwarz, Anne. 2004. Aspekte der Morphosyntax und Tonologie im Buli. Mit Schwerpunkt auf dem Buli von Wiaga. Dissertation, Humboldt-Universität zu Berlin.

-- (in preparation). Topic/Focus-Intersection in discourse: Distribution of copulative and narrative patterns in Gur.

Stassen, Leon. 1997. *Intransitive predication*: Oxford studies in typology and linguistic theory. Oxford u.a.: Clarendon Press.

Westermann, Diedrich. 1930. *A study of the Ewe language.* London: Oxford University Press.

Wilson, André A. 1963. Relative constructions in Dagbani. *Journal of African Languages* 2:139-144.

Abbreviations

CNJ	conjunction
COP	copula
DEF	definite (marker)
DEM	demonstrative (pronoun)
DEP	dependent (pronoun)
DISJ	disjunctive (pronoun)
dyn	dynamic (verb)
F	focus constituent
FM	focus marker
IND	indefinite (marker)
NC	narrative clause
NSF	non-subject focus (construction)
PROG	progressive marker
PRS	present (tense)

RC   RC

REL   relative (tense)

SF   subject focus (construction)

stat   stative (verb)

TMA   tense-mood-aspect

Ines Fiedler and Anne Schwarz
Humboldt-Universität zu Berlin
SFB 632 „Informationsstruktur"
Sitz: Mohrenstr. 40-41
Unter den Linden 6
10099 Berlin
Germany
sfb632.b1@rz.hu-berlin.de
www2.hu-berlin.de/gur_und_kwa_fokus

Diskurspragmatische Faktoren für Topikalität und Verbstellung in der ahd. Tatianübersetzung (9. Jh.)[*]

Roland Hinterhölzl, Svetlana Petrova und Michael Solf

Humboldt-Universität zu Berlin

The paper presents work in progress on the interaction between information structure and word order in Old High German based on data from the Tatian translation (9th century). The examination of the position of the finite verb in correspondence with the pragmatic status of discourse referents reveals an overall tendency for verb-initial order in thetic/all-focus sentences, whereas in categorical/topic-comment sentences verb-second placement with an initial topic constituent is preferred. This conclusion provides support for the hypothesis stated in Donhauser & Hinterhölzl (2003) that the finite verb form in Early Germanic serves to distinguish the information-structural domains of Topic and Focus. Finally, the investigation sheds light on the process of language change that led to the overall spread of verb-second in main clauses of modern German.

Keywords: information structure, verb placement in Germanic, language change

1   Zielsetzung

Der vorliegende Beitrag untersucht die Rolle der Informationsstruktur bei der Herausbildung der Wortstellungsregularitäten im heutigen Deutsch. Er ist ganz wesentlich von der Notwendigkeit bestimmt, eine Forschungslücke zu schließen − eine Forschungslücke, die ausgerechnet einer der wichtigsten und am intensivsten erforschten Zentralfragen der deutschen Syntax anhaftet, der Frage

---

nach der Herausbildung und Ausbreitung der Verbstellungsregularitäten im modernen Standarddeutschen.

Eine systematische und empirisch fundierte Untersuchung der Informationsstruktur in den früheren germanischen Sprachen steht in der bisherigen sprachhistorischen Forschung noch völlig aus. Zwar findet man sowohl in den klassischen junggrammatischen Untersuchungen von Braune (1894), Wackernagel (1891/92) und Behaghel (1932) als auch in der strukturalistischen Arbeit von Fourquet (1974) verstreute Hinweise darauf, dass informationsstrukturelle Kategorien Einfluss auf die Wortstellung der ältesten germanischen Sprachstufen nehmen[1], und ganz ähnliche Überlegungen werden auch in der Arbeit von Bean (1983) angestellt, die sich konkret mit der Entwicklung der Wortstellung im Englischen befasst. Diese Beobachtungen fließen jedoch nicht in ein kohärentes Gesamtbild zusammen, das Informationsstruktur innerhalb der Grammatik verortet und in die Erklärung von Sprachwandel einbezieht. Einen ersten Versuch in diese Richtung stellt der Ansatz von Kiparsky (1995) dar, der informationsstrukturelle Komponenten in ein Modell syntaktischen Wandels integriert.

Auffällig ist jedoch, dass die genannten Arbeiten insbesondere dann auf Faktoren der Informationsstrukturierung verweisen, wenn sie sich mit dem Problem der Verbstellung befassen. Bekanntlich zeichnet sich das finite Verb in den früheren Stufen der germ. Überlieferung durch eine hohe Positionsvielfalt im Satz aus und unterscheidet sich darin ganz wesentlich von der viel restringierteren Verbsyntax in den modernen Einzelsprachen. Dabei lässt die bisherige diachrone Syntaxforschung sowohl bei der Erklärung der Situation in der früheren Überlieferung als auch in der Frage nach der Herausbildung der

---

[1]   Einen ausführlichen Überblick über die bisherige Forschungssituation geben Donhauser & Hinterhölzl (2003).

heutigen Stellungseigenschaften des finiten Verbs einen klaren Zusammenhang zwischen Verbstellung und Informationsstruktur erkennen.

Eine Arbeit, die diesen Aspekt der historischen deutschen Syntax zuletzt weiter verfolgt hat, ist die Untersuchung zur linken Satzperipherie im ahd. Isidor bei Schlachter (2004). Die informationsstrukturelle Analyse eines kleinen Belegkorpus zeigt, dass die jeweils unterschiedliche Stellung des finiten Verbs (satzfinal vs. medial) in konjunktional eingeleiteten *dass*-Sätzen mit der inhaltlichen Hervorhebung von Teilen der Aussage korreliert. Schlachter zieht daraus den Schluss, dass das finite Verb in den untersuchten Sätzen des ahd. Isidors ein Mittel der informationsstrukturellen Gliederung darstellt, indem es die Fokusdomäne einer Aussage von anderen Teilen trennt, die den Hintergrund bilden.

Die Auffassung, dass das finite Verb im ältesten Deutsch als Mittel der Identifizierung informationsstrukturell relevanter Domänen fungiert und deshalb eine hohe Stellungsvarianz aufweist, wird durch Daten gestützt, die Donhauser & Hinterhölzl (2003) anderen Texten der ahd. Überlieferung stichprobenweise entnehmen. Dieser Befund verstärkt die Annahme, dass das finite Verb im früheren Deutsch Leistungen im Bereich der informationsstrukturellen Gliederung der Aussage übernimmt, d. h. an der Herstellung der Topik-Kommentar- bzw. Fokus-Hintergrund-Gliederung von Äußerungen im Diskurs beteiligt ist.

Das Ziel der vorliegenden Untersuchung ist es, diese Hypothese auf der Grundlage eines umfangreichen Belegkorpus aus der ahd. Überlieferung zu überprüfen. Eine solche Untersuchungsaufgabe erfordert zwei wesentliche methodische Vorüberlegungen. Zum einen setzt sie die Auswahl geeigneter ahd. Textzeugnisse voraus, die verlässliche Aussagen über die Interaktion von Verbstellung und Informationsstruktur an einer möglichst großen und für den

Originalzustand des Ahd. repräsentativen Datenmenge erlauben. Zum anderen muss die Methodik der informationsstrukturellen Analyse den Besonderheiten der historischen Korpusdaten angepasst werden. Da hier Informationen über die prosodische Gestaltung der Äußerungen fehlen, muss die Bestimmung informationsstruktureller Größen wie Topik und Fokus vor allem über die Einbeziehung pragmatischer Eigenschaften aus der Kontextanalyse erfolgen.

2   Methodische Vorüberlegungen

2.1   Textwahl und textspezifische Besonderheiten

Für die Zwecke der vorliegenden Untersuchung wurde der ahd. Tatian (St. Gallen Cod. 56) ausgewählt. Es handelt sich um eine Evangelienharmonie, die vom syrischen Kleriker Tatian bereits im 2. Jh. n. Ch. zusammengestellt und schon früh in andere Sprachen übersetzt worden ist. Auf der Grundlage einer lateinischen Version entstand um 830 in Fulda auch eine althochdeutsche (ostfränkische) Übersetzung, die von sechs in unterschiedlicher Intensität beteiligten Schreibern aufgezeichnet worden ist. Der heute in St. Gallen aufbewahrte Kodex überliefert den Text als Bilingue, in der althochdeutscher und lateinischer Text spaltenweise nebeneinander stehen, vgl. Faksimile-Seite (Anhang).

Ausschlaggebend für die Auswahl des ahd. Tatians sind dessen Umfang, Alter und Textsorte: Wir haben mit dem Tatian einen der umfangreichsten und ältesten Texte der ahd. Überlieferung vor uns, der darüber hinaus – im Gegensatz zu der ebenfalls umfangreichen Überlieferung des Evangelienbuchs von Otfrid von Weißenburg – in metrisch freier Form verfasst worden ist.

Eine besondere Herausforderung bei der Arbeit am ahd. Tatian stellt die Tatsache dar, dass es sich dabei um einen Übersetzungstext handelt, der

syntaktisch in hohem Maße die mitüberlieferte lateinische Vorlage nachahmt. Der hohe Grad an Originalabhängigkeit äußert sich in der wortgetreuen Wiedergabe der Vorlage über weite Textstrecken und in der Verwendung von syntaktischen Konstruktionen, die in den germanischen Sprachen nicht heimisch sind. Typisch für die lateinisch beeinflusste Lehnsyntax des ahd. Tatians ist etwa der Gebrauch des Dativus absolutus als Entsprechung zum lateinischen Ablativus absolutus, vgl. (1)

(1)   & ihesu / baptizato & orante / confestim ascendit de aqua,
         Inti themo heilante / gitoufitemo Inti b&ontemo /sliumo úfarsteig fon
         themo uuazzare. (ahd. T 48, 30–31)
         ‚Nachdem der Heiland getauft worden war, stieg er betend schnell aus
         dem Wasser'

Um die Zahl der aus einer derartigen Nähe zur Vorlage folgenden Fehlschlüsse möglichst gering zu halten, wird die Untersuchung zunächst grundsätzlich auf Differenzbelege beschränkt, auf diejenigen Stellen also, an denen die althochdeutsche Übersetzung von der Wortstellung der Vorlage signifikant abweicht. Solche Differenzbelege finden sich in der ahd. Tatian-Übersetzung erstaunlich oft. Ihr hoher Stellenwert für die Beschreibung der Tatian-Syntax wurde bereits in den frühesten Untersuchungen erkannt, vgl. Ruhfus (1897). In neuester Zeit werden Differenzbelege im ahd. Tatian in der systematischen Abhandlung von Dittmer & Dittmer (1997) gezielt zum Gegenstand der Untersuchung.

Von den zahlreichen Fällen syntaktischer Abweichungen sind für die vorliegende Untersuchung besonders solche Fälle interessant, die die relative Stellung von Konstituenten in Bezug auf das finite Verb betreffen. Dazu gehört die Änderung der im Original vorgegebenen Konstituentenabfolge ebenso wie die Hinzufügung und Auslassung von Konstituenten gegenüber dem Latein. Für

die Zwecke unserer Untersuchung weniger hilfreich sind die wesentlich zahlreicheren Umstellungen innerhalb von NPen (Umstellung von adjektivischem bzw. Genitivattribut und Nomen u. Ä.), weshalb sie vorerst aus der Datenerhebung ausgeschlossen werden. Im Allgemeinen macht die Orientierung auf Differenzbelege das Nebeneinander von fremdsprachigem Ausgangstext und Übersetzung für die Ermittlung von authentischem ahd. Material optimal nutzbar.

Eine weitere Besonderheit, die die Übersetzungstechnik des ahd. Tatians prägt und bei der Analyse seiner Syntax unbedingt Beachtung finden muss, liegt darin, dass jede Zeile des ahd. Textes ziemlich genau das Material in der entsprechenden lateinischen Zeile wiedergibt. Dieses Prinzip der zeilenweisen Übersetzung wird im Tatian konsequent beachtet und schränkt die Möglichkeit für Umstellungen der Konstituentenabfolge zum Original gravierend ein, vgl. Masser (1997 a und b). Eine neue, handschriftennahe Ausgabe des Textes, die Masser (1994) vorgelegt hat, bietet im Vergleich zu älteren Ausgaben die Möglichkeit, diesem wichtigen Übersetzungsprinzip bei der Behandlung syntaktischer Phänomene im ahd. Tatian Rechnung zu tragen.

2.2   Methodik der informationsstrukturellen Analyse

Ein grundlegendes Problem für die informationsstrukturelle Analyse an historischen Korpora stellt das weitgehende Fehlen direkter prosodischer Information dar. Deshalb müssen insbesondere die pragmatischen Eigenschaften der Konstituenten berücksichtigt werden, die für die Identifikation informationsstruktureller Grundparameter wie Topik und Fokus in Betracht kommen. Dazu gehört prototypisch der pragmatische Status von Diskursreferenten. In Hinblick darauf darf die Bekanntheit bzw. Vorerwähntheit von Information als Voraussetzung für Topikalität bzw. Topikfähigkeit von

Konstituenten gelten, andererseits steht die Neuheit bzw. Nicht-Vorerwähntheit von Satzmaterial mit informationsstruktureller Relevanz in Verbindung und gehört daher in die Domäne des Präsentationsfokus (‚Neuinformationsfokus') des Satzes.

Als Diskursreferenten werden in Anlehnung an Karttunen (1976) Individuen (Personen, Events, Objekte, Fakten etc.) aufgefasst, auf die in einem fortlaufenden Diskurs mit einem koreferenten Pronomen oder einer koreferenten definiten NP zurückverwiesen werden kann. Dies ist dann der Fall, wenn impliziert wird, dass die bezeichnete Entität in der textinternen Welt existiert und damit als Bezugspunkt in einem zusammenhängenden Diskurs zur Verfügung steht.

Über die Bestimmung des pragmatischen Status von Diskursreferenten liegt in der Literatur eine umfangreiche Diskussion vor, die eine über die Dichotomie von *gegeben* vs. *neu* hinausgehende Klassifikation begründet, vgl. Prince (1981). Unsere Untersuchung übernimmt das in Dik (^{2}1997, 315 ff.) vorgeschlagene Schema, das ausgehend vom informationellen Status der Diskursreferenten im jeweiligen Kontext folgende pragmatische Funktionen unterscheidet, die mit der Zuordnung informationsstruktureller Größen wie Topik und Fokus eng zusammenhängen:

- Ersteinführung von Diskursreferenten
- Anaphorische Referenz
- Nichtvorerwähnte, jedoch kontextuell erschließbare Referenten
- Wiederaufnahme von Diskursreferenten

Die implementierte Arbeitsmethodik sieht als ersten Schritt die Ermittlung des pragmatischen Status der Diskursreferenten aus der Kontextanalyse vor. Danach werden Beobachtungen über die Verwendung von Mitteln, die den Diskursstatus von Referenten markieren, angestellt. Vor allem interessiert dabei, ob in

Abhängigkeit vom pragmatischen Status von Diskursreferenten ein bestimmter Verbstellungstyp bevorzugt auftritt (V/1, V/2, V/end). Schließlich soll bei der Systematisierung der Beobachtungen über die Korrelation zwischen Diskursstatus und Verbstellung auch der Einfluss der Textstruktur auf die Wortstellung betrachtet werden.

3   Verbstellung und pragmatischer Status von Diskursreferenten im ahd. Tatian

3.1   Ersteinführung von Diskursreferenten

Die Untersuchung wendet sich zunächst Sätzen zu, in denen eine Entität erstmalig als Diskursreferent etabliert wird. Eine gesonderte Klasse von Sätzen, die ausschließlich auf die Funktion der Ersteinführung von Diskursreferenten spezialisiert ist, liegt im Fall der sog. Existentialkonstruktionen und Präsentationssätze (presentational sentences) vor. Diese zeichnen sich dadurch aus, dass sie typischerweise textinitial oder am Anfang eines neuen Erzählabschnitts auftreten, mit einer kleinen Gruppe nicht-agentivischer, zustandsbezeichnender Verblexeme vorkommen und dabei oft ein rahmensetzendes Adverbiale, d. h. eine Orts- oder Zeitangabe bzw. ein Expletivum (dt. *es* oder engl. *there*) in satzinitialer Position aufweisen.

Eine charakteristische Eigenschaft von Existentialkonstruktionen und Präsentationssätzen besteht darin, dass sie einen Kommentar über den in ihnen erwähnten Diskursreferenten nicht selbst abgeben, sondern ein solches im darauffolgenden Diskurs erst möglich machen. Prädikationsstrukturell betrachtet handelt es sich dabei also um Sätze ohne ein explizites Topik (Lambrecht 1994, 137–146, 177–181) bzw. um Sätze, in denen die Durchführung der Topik-Kommentar-Gliederung vollständig unterbleibt (Drubig 1992). Durch seine

Neuheit im Diskurs ist der erwähnte Referent Bestandteil der Domäne des Neuinformationsfokus, die in diesem Fall den ganzen Satz umfasst (all-new bzw. all-focus-Sätze). Präsentationssätze haben demnach – wie textinitiale Sätze im Allgemeinen auch – eine thetische Lesart.

Im ahd. Tatian zeichnet sich in Existentialkonstruktionen und Präsentationssätzen eine deutliche Präferenz für die Anfangsstellung des finiten Verbs ab. Die Konstituente, die den einzuführenden Diskursreferenten bezeichnet, folgt in diesem Fall dem finiten Verb nach.

In einem Teil der Belege liegt die Anfangsstellung des Verbs bei Ersteinführung von Diskursreferenten bereits im Lateinischen vor. Bemerkenswert ist jedoch, dass das Ahd. eine klare Tendenz dazu aufweist, die im Latein vorliegende V/1-Struktur bei Ersteinführung von Diskursreferenten aufrechtzuerhalten, vgl. (2)–(3):

(2)   **Fuit** in diebus herodis regis / iudeę <u>quidam sacerdos</u> / […]/ & <u>uxor illi</u>
         uuar [sic!] In tagun herodes thes cuninges / Iudeno <u>sumer biscof</u> / […]/
         Inti <u>quena Imo</u> (ahd. T 25, 29 ff.)
         ‚Es war zu der Herrschaftszeit von König Herodes von Judäa ein Priester
         […] und dessen Frau'

(3)   **Erat** autem <u>quidam languens</u> / lazarus a b&hania
         uuas <u>sum siocher</u> / lazarus fon b&haniu (ahd. T 228, 27 f.)
         ‚Es war ein kranker Mann, Lazarus von Bethanien.'

Noch stärker fällt jedoch ins Gewicht, dass die Anfangsstellung des Verbs bei Ersteinführung von Diskursreferenten auch in Differenzbelegen, d. h. nur in den ahd. Sätzen und gegen die Wortstellung der Vorlage, vorzufinden ist. In (4) und (5) wird eine mediale Verbstellung der lat. Vorlage zugunsten einer V/1-Stellung im Ahd. verändert, die neu einzuführenden Diskursreferenten *hirta*

‚Hirten' und *uuitua* ‚eine Witwe' erscheinen in den ahd. Belegen gegen das Original nach dem finiten Verb:

(4)   Et <u>pastores</u> **erant** In regione eadem
         uuarun thô <u>hirta</u> In thero lantskeffi[2] (ahd T 35, 29)
         ‚Da waren Hirten in jener Gegend'

(5)   <u>Vidua</u> autem quaedam **erat** / In ciuitate illa
         uuas thar ouh sum <u>uuitua</u> / In thero burgi (ahd. T 201, 2)
         ‚Es war dort auch eine Witwe in dieser Stadt'

Die Struktur solcher Sätze kann durch die in (6) angegebene Formel repräsentiert werden:

(6)   Ersteinführung von Diskursreferenten in Präsentationssätzen

         V/1: $_{FOC}$[Vfin … DR$_{neu}$ …]

Neben den Fällen von verbinitialer Stellung bei Ersteinführung von Diskursreferenten finden sich in derselben Kontextfunktion auch V/2-Sätze, die links vor dem finiten Verb ein rahmensetzendes Adverbiale aufweisen, vgl. die Zeit- bzw. Ortsadverbiale ahd. *thô* ‚da, damals' in (7) und *thar* ‚dort' in (8):

(7)   & ecce <u>homo</u> **erat** In hierusalem.'
         senonu tho **uuas** <u>man</u> In hierusalem.' (ahd. T 37, 23)
         ‚Und siehe, da war ein Mann in Jerusalem'

(8)   **erant** autem ibi <u>lapideę hydrię</u>
         thar **uuarun** <u>steininu uuazzarfaz</u> (ahd. T 81, 26)
         ‚Dort waren steinerne Wasserfässer'

---

[2] Das finite Verb wird fett angegeben, der Diskursreferent, der Gegenstand der Analyse ist, wird unterstrichen.

Fälle der Ersteinführung von Diskursreferenten, die in V/2-Strukturen mit vorangestellten Frame-Adverbialen realisiert werden, können unter der in (9) angegebenen Formel systematisiert werden:

(9)   Ersteinführung von Diskursreferenten

Frame + V/2   *tho*/*thar* $_{FOC}$[Vfin … DR$_{neu}$ ...]

Beim Vergleich von (6) und (9) wird deutlich, dass die Beispiele von V/2-Stellungen mit den in (2)−(5) aufgezeigten V/1-Sätzen vor allem eines gemeinsam haben: Die Konstituente, die dem neu eingeführten Diskursreferenten entspricht, erscheint in beiden Strukturtypen gleichermaßen rechts vom finiten Verb. Die postverbale Stellung von Ausdrücken, die neue Diskursreferenten etablieren, wird in V/2-Strukturen des Ahd. auch gegen das Latein erzeugt. Das wird in Fällen wie (7) besonders deutlich, wo die Stellung von ahd. *man* ‚ein Mann' dem finiten Verb gegen das Original nachfolgt.

Für eine Dominanz der V/1-Stellung über die *thô*/*thar*-V/2-Stellung bei Ersteinführung von Diskursreferenten spricht die Tatsache, dass sich Belege finden lassen, in denen ein Zeit- bzw. Ortsadverbiale gegen die Originalvorlage eingefügt wird, und zwar nach dem finiten Verb, vgl. (10)–(11):

(10)   & **erat** <u>anna proph&issa</u>
        uuas thô thâr <u>anna uuizzaga</u> (ahd T 38, 22)
        ‚Es war damals dort die Prophetin Anna'

(11)   & **erat** <u>quidam regulus</u>
        uuas thar <u>súm rihtari</u> (ahd. T 90, 10)
        ‚Es war dort ein Richter'

Dieselbe postverbale Stellung neuer Diskursreferenten begegnet ferner in Sätzen, in denen sie als Objekte transitiver Verben realisiert werden, vgl. (12):

(12)   Et praeteriens **uidit** <u>hominem</u> cecum / a nauitate [sic!]
        furfarentj **gisah** <u>man</u> blintan / fon giburtj (ahd. T 220, 13 f.)
        ‚Vorbeigehend sah er einen Mann, der von Geburt an blind war.'

Schließlich lassen sich bei Ersteinführung von Diskursreferenten im ahd. Tatian gelegentlich auch weitere Abfolgemuster beobachten; diese können jedoch als Ausnahmen erklärt werden. Sie sind entweder durch die Einhaltung des Zeilenprinzips wie in (13) bedingt oder stellen Nachbildungen des Originals dar, dazu (14):

(13)   ecce <u>defunctus</u> / **efferebatur**
        senu <u>arstorbaner</u> / **uúas** gitragan (ahd. T 84, 22 f.)
        ‚Siehe, es wurde ein Toter getragen'

(14)   <u>multae uiduae</u> **erant** / in diebus heliae in israhel
        <u>manago uuituuuvn</u> **uuarun**/ in heliases tagon in israhel (ahd. T 114, 27f.)
        ‚Es waren viele Witwen in den Tagen Elias' in Israel'

3.2   Anaphorische Referenz

In diesem Abschnitt soll die Situation in Sätzen betrachtet werden, in denen eine Aussage über bereits etablierte Diskursreferenten getroffen wird. Hier unterscheiden wir zwischen Fällen der linearen thematischen Progression bei einem singulären Diskursreferenten (lineare thematische Progression) und dem sog. Topikwechsel, d. h. dem selektiven oder kontrastiven Wechsel zwischen mehreren bereits eingeführten Diskursreferenten.

3.2.1 Lineare thematische Progression

In dieser Kategorie sind Fälle versammelt, in denen ein bereits eingeführter Diskursreferent zum Gegenstand einer unmittelbar nachfolgenden Aussage wird. In den Begrifflichkeiten der Prager Schule liegt hier der klassische Fall der linearen thematischen Progression vor: Das Rhema des vorangehenden Satzes wird zum Thema des Folgesatzes. Die Weiterführung des Diskursreferenten erfolgt durch pronominale Wiederaufnahme oder durch Verwendung einer koreferenten definiten NP.

Anders als bei der Ersteinführung von Diskursreferenten, bei der eine klare Präferenz für die V/1-Stellung sowie für die postverbale Realisierung der jeweiligen Diskursreferenten beobachtet wurde, sind bei der unmittelbaren Weiterführung bereits bekannter Referenten bei einem Großteil der Differenzbelege V/2-Strukturen zu finden. Dabei wird die Stelle links vor dem finiten Verb durch den bereits vorerwähnten Diskursreferenten selbst besetzt. Diese Konstellation ist in Minimalpaaren von aufeinanderfolgenden Sätzen zu beobachten, in denen der erste Satz eine V/1-Stellung bei Ersteinführung eines Diskursreferenten aufweist und der zweite Satz denselben Diskursreferenten in die satzinitiale Stelle eines V/2-Satzes übernimmt, vgl. (15):

(15) **Fuit** […] <u>quidam sacerdos</u> / […] / & <u>uxor illi</u> […] / **erant** autem iusti
      ambo ante deum
      uuar [sic!] […] <u>sumer biscof</u> /[…] / Inti <u>quena Imo</u> […] / <u>siu</u> **uuarun**
      rehtiu beidu fora gote (ahd. T 25, 29–26, 3)
      ‚Es war […] ein Bischof […] und dessen eine Frau […]. Sie waren beide
      rechtschaffen vor Gott.'

Die Prädikationsstruktur dieser Sätze erlaubt es, sie als kategorische Sätze mit einer Topik-Kommentar-Gliederung zu interpretieren. Die Konstituente, die im Anschlusssatz links vom finiten Verb steht, ist ihrem pragmatischen Status nach

Topik im Sinne der familiarity- und aboutness-Konzepte (vgl. die Zusammenstellung in Frey 2000: 137 f.). Das finite Verb eröffnet die Domäne, in der ein Kommentar zum Satztopik erfolgt, und markiert gleichzeitig den Beginn der Domäne, die neue Information im Diskurs präsentiert.

Die Struktur, die sich damit für Sätze mit anaphorischer Referenz im Ahd. ergibt, kann folgendermaßen dargestellt werden:

(16)   Anaphorische Referenz

V/2 $_{TOP=BGR}[DR_{giv}]$ $_{FOC=COMMENT}[Vfin ...]$

Dieses Abfolgemuster in kategorischen Sätzen mit bekannten Diskursreferenten wird im Ahd. auf vielfältige Weise gegen das Original konstruiert. Zunächst werden Belege angeführt, in denen die Übersetzung zwar die im Original vorliegenden Konstituenten übernimmt, ihre Abfolge jedoch konsequent in Richtung der in (16) ausgewiesenen Struktur umstellt:

(17)   ego sum pastor bonus. bonus pastor / animam suam **dat** pro ouibus suis
         ih bin guot hirti. guot hirti / **tuot** sina sela furi siniu scaph.
         (ahd. T 225, 16 f.)
         ,Ich bin ein guter Hirte. Der gute Hirte gibt seine Seele für seine Schafe.'

(18)   ecce defunctus / efferebatur. filius unicus / matris suae. & haec uidua
         erat.
         senu arstorbaner / uúas gitragan einag sun / sinero muoter Inti thiu **uuas**
         uuituuua (ahd. T 84, 22 ff.)
         ,Siehe, da wurde ein Toter getragen, der einzige Sohn seiner Mutter, und
         diese war Witwe'

In diesen Belegen wird die Spät- bzw. Endstellung des Originals zugunsten einer V/2-Stellung im Ahd. verändert, wobei die einzige Stelle vor dem finiten Verb der unmittelbar vorerwähnten Konstituente als Topik der Aussage vorbehalten

bleibt. Der Kommentar darüber bzw. die Domäne der Neuinformation wird mit dem finiten Verb eingeleitet.

Ferner wird im Ahd. die in (16) angegebene V/2-Struktur bei anaphorischer Referenz auch durch Einsetzung von im Original fehlenden Konstituenten erzeugt. Dazu gehört die Hinzufügung der finiten Kopula in Zweitstellung bei der Wiedergabe elliptischer Kopulativkonstruktionen des Lateins oder bei Auflösung synthetischer Verbformen wie in (19) bzw. die Einsetzung des Subjektpronomens links vom finiten Verb wie in (20):

(19)   lazarus **Infirmabatur**
        lazarus **uuard** cumig (ahd. T 229, 3)
        ‚Lazarus ist krank geworden'

(20)   & uxor tua […] / pari& tibi <u>filium</u>. / […] / & **erit** tibi gaudium &
        exultatio […] / **erit** enim magnus coram domino
        Inti thin quena […] / gibirit thir <u>sun</u>. / […] / Inti <u>her</u> **ist** thir gifeho Inti
        blidida […]/ <u>her</u> **ist** uuârlihho mihhil fora truhtine (ahd. T 26, 25 ff.)
        ‚Und deine Frau […] wird dir einen Sohn gebären. […] Er wird dir
        Freude und Wonne sein. [...] Er wird wahrlich groß vor Gott sein'

Bei einem Großteil der ermittelten Belege kommen die Umstellung und Hinzufügung von Konstituenten gegen das Latein kombiniert zur Anwendung, vgl. (21):

(21)   [Auf die Frage hin: „Wie sind dir deine Augen wieder aufgetan
        worden?":]
        <u>Ille homo</u> qui dicitur / ihesus. lutum **fecit** & unxit / oculos meos
        <u>ther man</u> thiedar ist giquetan / heilant. <u>her</u> **t&a** leimon inti salbota / minu
        ougun (ahd. T 221, 9 f.)
        ‚Dieser Mann, der da genannt wird Heiland, er rührte Schlamm ein und
        salbte [damit] meine Augen'

Starke Evidenz dafür, dass die in (16) aufgezeigte V/2-Struktur mit der pragmatischen Funktion der anaphorischen Referenz zusammenhängt, liefern Differenzbelege in eingeleiteten Nebensatzstrukturen des Ahd. Unmittelbar vorerwähnte Konstituenten, die im Original postverbal erscheinen, werden im Ahd. systematisch in die Position links vor dem finiten Verb gestellt:

(22)   & non erat <u>illis</u> filius. eo quod / **ess&** elisab&h sterilis
         Inti niuuard <u>In</u> sun. bithiu uuanta / elisab& **uuas** unberenti (T 26, 6 f.)
         ‚Und sie hatten keinen Sohn, weil Elisabeth unfruchtbar war‘

(23)   Lucerna corporis. est <u>oculus</u>. / si **fuerit** <u>oculus tuus</u> simplex. / totum
         corpus tuum lucidum **erit**.
         liohtfaz thes lihhamen ist <u>ouga</u> / oba <u>thin ouga</u> **uuirdit** luttar / thanne **ist**
         al thin lihhamo liohter (ahd. T 69, 21 ff.)
         ‚Das Licht des Körpers ist das Auge. Wenn dein Auge hell wird, dann
         wird auch dein ganzer Körper licht sein‘

Dieselbe Art der Konstituentenumstellung wird in eingeleiteten Nebensätzen des Ahd. selbst dann vorgenommen, wenn der entsprechende lat. Satz die V/end-Stellung aufweist, vgl. (24)−(25):

(24)   [...] ueni / ut […] <u>qui uident</u> caeci **fiant**
         [...] quam ih […] / thaz [...] <u>thie dar gisehent</u> daz <u>sie</u> **sin** blinte
         (ahd. T 224,4 ff.)
         ‚Ich kam […] damit diejenigen, die sehen, blind werden‘

(25)   domine ego credidi / quia tu es christus filius dei / <u>qui</u> In mundum
         uenisti
         trohtin ih giloubta. / thaz thu bist crist gotes sun. / <u>thie dar</u> **quam** In
         mittilgart (T 231, 18 ff.)
         ‚Herr, ich glaubte, dass du Christus, der Sohn Gottes bist, der auf die Welt
          kam = und dieser kam auf die Welt‘

Sowohl vor der Maßgabe der maximalen Anlehnung an die Originalwortfolge in diesem Text als auch vor dem Hintergrund der Annahme, dass die V/end-

Stellung im eingeleiteten Nebensatz des Ahd. – so wie im heutigen Deutsch auch – die kanonische Verbstellung repräsentiert (Lenerz 1984), ist die Evidenz in (24)−(25) überaus erstaunlich. Für die wissenschaftliche Auswertung dieses Textes sowie für die verfolgte Argumentationslinie sind diese Belege in zweierlei Hinischt besonders von Belang. Zum einen relativieren sie die von Dittmer & Dittmer (1998, 18) aufgestellte Behauptung, dass der ahd. Tatian bezüglich seiner Vorlage an keiner Stelle in Richtung einer für das heutige Deutsch untypischen Syntax abweiche. Zum zweiten stehen sie pragmatisch und strukturell vollkommen in Einklang mit den anderen vorab präsentierten Fällen einer Weiterführung bekannter Diskursreferenten: Es ist auffällig, dass die Konstituente links vom finiten Verb auch hier vorerwähnt ist, während die Domäne nach dem finiten Verb das inhaltliche Hauptgewicht der Mitteilung trägt.

Diese Daten bestätigen die Beobachtung, dass V/2-Stellungen im Ahd. regelmäßig mit der pragmatischen Funktion der anaphorischen Referenz korrelieren, und stützen weitgehend die Hypothese, dass die Verbstellung im Ahd. als Mittel der informationsstrukturellen Gliederung der Aussage eingesetzt wird.

3.2.2 Topikwechsel

Oft bezieht sich der Satz nicht auf einen einzigen, sondern auf mehrere Diskursreferenten, die vorerwähnt und daher im Gedächtnis der Diskurspartizipanten aktiviert, d. h. salient sind. Es stehen damit mehrere potentiell geeignete Topik-Kandidaten im Diskurs bereit. Von entscheidender Bedeutung ist dabei die Tatsache, dass diesen Diskursreferenten gemäß ihrer syntaktischen Realisierung im vorangehenden Diskurs ein unterschiedlicher Grad an Salienz zukommt, der für die Gestaltung des Nachfolgesatzes von Bedeutung ist. Ein gängiges Modell, die Salienz von Diskursreferenten auf einer

Skala abzubilden, liegt in der Abstufung Subjekt > direktes Objekt > indirektes Objekt > Rest vor. Dabei nimmt man an, dass der in der vorangehenden Äußerung salienteste Diskursreferent seinen Status auch in der nachfolgenden Struktur behält, wenn ein Wechsel in der Salienzskala nicht durch besondere Mittel angezeigt wird. Zu diesen Mitteln zählen etwa die Genuskongruenz, die Wahl von Anaphern mit einem höheren Grad an sprachlicher Explizitheit (Ø-Anapher, Personalpronomen, volle NP als Eigenname, nominaler Ausdruck etc.) bzw. der Unterschied zwischen Personal- und Demonstrativpronomen im Deutschen, vgl. Lambrecht (1994, 204 f.). Auf diese Weise ist dafür gesorgt, dass der jeweilige anaphorische Ausdruck mit dem richtigen Antezedens identifiziert und die Äußerung angemessen interpretiert werden kann.

Die vorliegenden Beobachtungen konzentrieren sich auf Strategien zur Kennzeichnung eines Wechsels in der Salienzskala von anaphorischen Ausdrücken, wobei grundsätzlich zwischen einem selektiven Wechsel von Diskursreferenten und einem zusätzlich mit Merkmalen der kontrastiven Gegenüberstellung verknüpften Wechsel unterschieden wird.

(+ contrast / + selective)

Bei Ausdrücken mit diesen Merkmalen handelt es sich um kontrastive Topiks. In einem Teil der Belege wird diese pragmatische Funktion mit der für die anaphorische Referenz typischen V/2-Stellung auch gegen die Wortstellung in der Vorlage realisiert. Der Kontrast wird möglicherweise auch durch prosodische Hervorhebung der in Kontrast gesetzten Konstituenten erzielt, wofür sich in einem historischen Korpus jedoch kaum sichere Angaben ermitteln lassen. In (26) wird allerdings eine Kontrastbetonung der Konstituenten ziemlich sicher durch die Setzung der deiktischen Personalpronomen lat. *tu* ‚du' und lat. *nos* ‚wir' in Subjektstellung nahegelegt.

(26)   tu discipulus illius **sis** / <u>Nos</u> autem moysi discipuli **sumus**
        <u>thu</u> **sís** sín iungiro / <u>uuir</u> **birumes** moyseses iungiron (ahd. T 223, 7 f.)
        ‚Du mögest sein Jünger sein, wir sind Jünger des Moses'

Während in (26) der eigentliche Kontrastmarker – die Partikel lat. *autem* ‚aber'
– nicht mitübersetzt wurde, findet sich in anderen Belegen mit kontrastiver
Lesart eine entsprechende Markierung durch Partikel bzw. Adverbien.

Eine interessante Beobachtungsmöglichkeit bietet die ahd. Übersetzung
der kontrastiven Partikel lat. *autem* in der häufig vorkommenden Wendung lat.
ego autem dico ‚ich aber sage'. Die Wiedergabe von *autem* durch *thanne* im
Ahd. hängt in diesen Fällen ziemlich regelmäßig mit dem Vorliegen einer
kontextuell deutlich erschließbaren Kontrastivität zusammen. Fest steht, dass der
ahd. Text das Adverb *thanne* als Entsprechung zu lat. *autem* stets dort enthält,
wo sich der Sprecher (Jesus) in irgendeiner Weise inhaltlich von den zitierten
Aussagen der früheren Propheten distanziert. Das betrifft etwa die Auffassung
vom Schwören in (27), von der persönlichen Rache in (28) und von der
Scheidung in (29):

(27)   Dictum est autem.' Quicumque dimiserit. / uxorem suam.' d& illi
        libellum repudii; / <u>Ego autem</u> **dico** uobis., / quia omnis qui dimiserit /
        uxorem suam […] / facit eam moechari.
        íz ist giq&an só uuér so fúrlaze. / sina quenun. gebe iru buoh thanatribes.
        / <u>thanne ih</u> **quidu** íu / thaz thero giuuelih thie furlazit / sina quenun […] /
        tuot sia furligan (ahd. T 64, 6 ff.)
        ‚Es ist auch gesagt: „Wer sich von seiner Frau scheidet, der soll ihr einen
        Scheidebrief geben." Ich aber sage euch: „Wer sich von seiner Frau
        scheidet, […], der macht, dass sie die Ehe bricht" […]'

(28)   Audistis quia dictum est antiquis. / non perierabis […] <u>ego autem</u>
       dico uobis., / non iurare omnino
       Ir gihórtunt thaz then alton giq&an uúas / nifúrsuueri thih. […] <u>thanne</u>
       <u>ih</u> **quidu** íu / thaz mán zi thuruhslahti nisuuere (ahd. T 64, 13 ff.)
       ‚Ihr hörtet, dass zu den Alten gesagt worden ist: „Du sollst nicht falsch
       schwören!" Ich aber sage euch, dass man überhaupt nicht
       schwören soll.'

(29)   Audistis quia dictum est / oculum pro oculo. [...] <u>ego autem</u> **dico** uobis.,
       / non resistere malo
       Ir gihórtut thaz giqu&an ist / ouga furi ouga. [...] / <u>thanne ih</u> **quidu** íu
       / thaz ír niuuidarstant& ubile. (ahd. T 64, 29 ff.)
       ‚Ihr hörtet, dass gesagt worden ist: „Auge um Auge." Ich aber sage euch,
       dass ihr dem Übel nicht widerstreben sollt.'

Bei inhaltlicher Übereinstimmung mit der Aussage der Propheten, so etwa
bezüglich der Auffassung von Totschlag (30) und Ehebruch (31), bleibt lat.
autem im Ahd. unübersetzt:

(30)   Audistis quia dictum est. / antiquis.' non occides. […] / <u>ego autem</u> **dico**
       uobis. / quia omnis qui irascitur / fratri suo. reus erit iudicio
       ír gihórtut thaz giqu&an uúas. / then alton. niuúis manslago / [...] <u>ih</u>
       quidu íu / thaz iogiuuelih ther sih gibilgit / zi sinemo bruoder. ther ist
       sculdig duomes (ahd. T 62, 21 ff.)
       ‚Ihr habt gehört, dass zu den Alten gesagt wurde: „Du sollst nicht töten."
       Ich sage euch, dass, wer seinem Bruder zürnt, des Gerichts schuldig ist.'

(31)   Audistis quia dictum est / antiquis.' non moechaberis., / <u>Ego autem</u> **dico**
       uobis., / quoniam omnis qui uiderit mulierem / ad concupiscendum eam.'
       / Iam moechatus est eam in corde suo.
       Ir gihortut thaz giqu&an ist / then alton nifurligi thíh / <u>ih</u> **quidu** íu / thaz
       iogiuuelih thiethar gisihit uúib / sie zigeronne / iu habet sia forlegana in
       sinemo herzen (ahd. T 63, 18 ff.)
       ‚Ihr hörtet, dass zu den Alten gesagt wurde: „Du sollst nicht
       ehebrechen!" Ich sage euch, dass wer eine Frau ansieht, um sie zu
       begehren, mit ihr schon die Ehe gebrochen hat in seinem Herzen'

Der syntaktische Status dieser Kontrastmarkierung ahd. *thanne* ist zweifelhaft. Sie kann sowohl als Modifizierer innerhalb der NP als auch als selbständige syntaktische Konstituente mit Satzgliedstatus aufgefasst werden. Deshalb lässt sich auch die Verbstellung in solchen Fällen (V/2 vs. V3) nicht eindeutig bestimmen.

(- contrast; + selective)

Selektiver Topikwechsel wird im ahd. Tatian ebenfalls durch Partikeln oder Adverbien signalisiert, wobei zugleich auch eine Tendenz zur Früherstellung des finiten Verbs zu beobachten ist, vgl. (32) und (33). Der V/2-Status der Sätze ist allerdings weitgehend unklar, vgl. besonders (33):

(32)   (Nach dem Erscheinen des Engels, während Zacharias das Rauchopfer bringt:)
       & zacharias turbatus **est**
       thanan tho <u>zacharias</u> **uuard** gitruobit (ahd. T 26, 20)
       ‚Zacharias aber erschrak davor'

(33)   **mansit** autem <u>maria</u> cum illa / quasi mensibus tribus & reuersa **est** / In domum suam, / <u>Elisab&h</u> autem Impl&um est / tempus pariendi
       uuon&a <u>maria</u> mit Iru / nah thri manoda Inti **uuarb** / zi Ira hûs, / <u>Elisab&h</u> uuârlihho **uuard** gifullit / zît ziberanne (ahd. T 30, 13 ff.)
       ‚Maria blieb bei ihr [Elisabeth] etwa drei Monate lang und ging zurück in ihr Haus. Für Elisabeth aber erfüllte sich die Zeit zu gebären'

Von einer Signalfunktion der Partikeln und Adverbien bei einer Verschiebung der Salienz von Diskursreferenten kann auch in Fällen häufiger lexikalischer Wiederholungen ausgegangen werden, die die Herstellung der anaphorischen Beziehung gefährden, vgl. (34):

(34)   Quis uestrum hab& <u>amicum</u> / & ibit ad <u>illum</u> media nocte / & dicit <u>illi</u>.
       ‚<u>Amice.</u>' / commoda mihi tres panes / quoniam amicus meus uenit de uia
       / ad mé [...] <u>& ille</u> de intus **dicat** […]
       Vuelih íuuer hab& <u>friunt</u> / Inti ferit zi <u>Imo</u> In mittero naht / Inti quidit
       <u>imo. friunt</u> / Intlih mír thriu brót / uuanta mín friunt quam fon uúege zi
       mir. […] <u>her thanne</u> fon innana **quede** […] (ahd T 72, 13 ff.)
       ‚Wer$_i$ von euch hat einen Freund$_j$, und geht zu ihm$_j$ mitten in der Nacht
       und sagt ihm$_j$: „Freund$_j$, gib mir$_i$ drei Brote, denn mein Freund$_k$ ist von
       der Reise zu mir$_i$ gekommen [...]" und er$_j$ würde dann von drinnen sagen:
       […]'

3.3   Nichtvorerwähnte, jedoch erschließbare Diskursreferenten

In diesem Teil betrachten wir Sätze mit Diskursreferenten, die im
vorangehenden Kontext zwar nicht explizit erwähnt sind, aber in einer
Analogie- oder Teil-Menge-Beziehung zu einem anderen, vorab eingeführten
Diskursreferenten stehen. Solche Diskursreferenten sind kontextuell bzw. aus
dem Weltwissen des Sprechers inferierbar und stellen geeignete Topik-
Kandidaten bereit, über die im jeweiligen Kontextbezug eine Aussage erwartbar
ist, vgl. (vgl. Prince 1981; Dik [2] 1997, 323 f.).

       In diesen Fällen ist im ahd. Tatian – wie bei den Fällen der anaphorischen
Referenz – eine Präferenz für V/2-Stellungen zu beobachten, wobei inferentiell
erschließbare Diskursreferenten in der für kontextuell vorerwähnte Topiks
typischen Stelle unmittelbar vor dem finiten Verb vorkommen:

(35)   & <u>nomen eius</u> elisab&h
       Inti <u>ira namo</u> **uuas** elisab&h (ahd. T 26,2)
       ‚Und ihr Name war Elisabeth'

In (35) ist die Konstituente ahd. *ira namo* ‚ihr Name' auf die vorab eingeführte
Person ahd. *quena* ‚[seine] Frau' bezogen. Die Nennung des Namens gilt nach
der Einführung der Handlungspersonen einer Erzählung durchaus als erwartbar.

Das gegen das Latein eingesetzte finite Kopulaverb steht hier zwischen der durch Inferenzbeziehung topikal zu deutenden Konstituente und der neuen Information im Prädikatsnomen und fungiert somit als trennendes Glied zwischen dem Topik und dem Neuinformationsfokus der Aussage. Solche Vorkommen fallen informationsstrukturell und syntaktisch mit den Fällen der Weiterführung bekannter Diskursreferenten zusammen und weisen dabei die Struktur von (16) auf.

3.4   Wiederaufnahme von Diskursreferenten

In diesem Teil werden Sätze betrachtet, in denen bekannte Diskursreferenten nach einer gewissen Unterbrechung in den Diskurs wiederaufgenommen werden.

Aufgrund des Umstands, dass sich Eigenschaften der Wiederaktivierung mit Merkmalen anderer pragmatischer Funktionen überschneiden (Ersteinführung, Wechsel von Diskursreferenten), ist es durchaus problematisch, der Wiederaufnahme den Status einer eigenständigen pragmatischen Klasse zuzuweisen. Darum überrascht auch nicht, dass die ahd. Sätze, in denen die Rede nach einer Unterbrechung erneut auf einen vorab erwähnten Diskursreferenten kommt, keine einheitliche syntaktische Struktur erkennen lassen. Zum einen finden sich hier Fälle von V/1-Stellungen in Korrespondenz-(36) aber auch in Differenzbelegen (37), wobei der wiedereingeführte Referent spät im Satz, rechts vom finiten Verb erscheint:

(36)   **erat** autem & <u>iohannes</u> baptizans
       Vuas ouh tho <u>iohannes</u> toufenti (ahd. T 56, 23)
       ‚Auch Johannes taufte dort'

(37)   & ecce <u>angelus domini</u>
         quam thara <u>gotes engil</u> (ahd. T 35, 32)
         ‚Da kam Gottes Engel'

In anderen Fällen, in denen die Wiedereinführung von Diskursreferenten begegnet, treten allerdings auch V/2-Sätze auf, wobei der wiedereingeführte Diskursreferent die Stellung vor dem finiten Verb einnimmt:

(38)   <u>puer autem</u> **crescebat** & confortabatur / spiritu
         <u>ther kneht</u> **uuvohs** Inti uuard gistrengisot / geiste (T 32, 6 f.)
         ‚Dieser Knabe wuchs und wurde gestärkt im Geiste'

Hier wird durch die Hinzufügung des einfachen Demonstrativpronomens ahd. *ther* ‚dieser' gegen das Latein die Bekanntheit bzw. Identifizierbarkeit des Diskursreferenten suggeriert, was eine Topik-Lesart der entsprechenden Konstituente nahe legt.

Darüber hinaus werden Partikeln bzw. Adverbien verwendet, die ähnlich wie im Fall des Wechsels aktiver Diskursreferenten eine Verschiebung auf der Salienzskala signalisieren; syntaktisch ist die Verbstellung nicht eindeutig als V/2 zu bestimmen:

(39)   <u>Puer autem</u> **crescebat**
         <u>ther kneht</u> uuârlihho **uuvohs** (ahd. T 42, 8)
         ‚Dieser Knabe wuchs wahrlich'

3.5   Zwischenbilanz

Die vorangehende Untersuchung hat gezeigt, dass sich bestimmte Verbstellungstypen im Tatian in der Regel mit bestimmten pragmatischen Funktionen von Diskursreferenten in Verbindung bringen lassen. So findet man bei der Ersteinführung von Diskursreferenten, etwa in Präsentationssätzen und Existentialkonstruktionen, vorzugsweise die V/1-Stellung. Dagegen begegnet

die V/2-Stellung regelmäßig in Sätzen mit kontextuell vorerwähnten bzw. erschließbaren Referenten, die in der Position vor dem finiten Verb stehen. Damit bestätigt sich am Beispiel des Tatians die Annnahme, dass zwischen der Stellung des finiten Verbs und dem pragmatischem Status von Diskursreferenten ein Zusammenhang besteht.

Wenn man dieses Fazit auf die Ebene der Informationsstruktur der betreffenden Sätze überträgt, lassen sich die Aussagen über die Verteilung der Verbstellungsmuster in Abhängigkeit vom informationellen Status von Diskursreferenten weiter präzisieren. Die Existenzialkonstruktionen und Präsentationssätze, die sich im Ahd. durch eine hohe Präferenz für V/1-Abfolgen auszeichnen, gelten als all-focus-Sätze ohne Satztopik. V/2-Stellungen sind wiederum typisch für Sätze mit Diskursreferenten, die sowohl im Sinne des familiarity- als auch im Sinne des aboutness-Konzepts als Satztopiks identifizierbar sind. In diesem Fall nimmt der bekannte Diskursreferent, d. h. der Topikausdruck, die Stellung unmittelbar vor dem finiten Verb ein.

Damit steht auf der Basis einer breiteren empirischen Erhebung fest, dass V/1- und V/2-Stellungen im Ahd. primär eine funktionale Opposition bilden, die im Rahmen der informationsstrukturellen Gliederung der Äußerung angesiedelt ist. Die Früherstellung des finiten Verbs im Ahd. gegenüber der V/end-Stellung, die als basisgeneriert angenommen wird, lässt sich mit Leistungen im Bereich der informationstrukturellen Gliederung der Aussage identifizieren: Genauer gesagt besetzt das finite Verb im Ahd. sowohl in V/1- als auch in V/2-Sätzen dieselbe Position, nämlich die Position, die den Beginnn der Domäne des Neuinformationsfokus auszeichnet. Der entscheidende Unterschied zwischen den Verbstellungstypen besteht darin, dass V/1-Strukturen all-focus-Sätze ohne ein explizites Topik darstellen, während V/2-Sätze mit satzinitialer Topik-Konstituente als in Topik und Kommentar bzw. Fokus und Hintergrund

gegliedert erscheinen, wobei das finite Verb hier die Domänen von Topik und Kommentar/Fokus voneinander abgrenzt.

Den ausgewiesenen Früherstellungen – V/1 und V/2 im Ahd. – kann die folgende gemeinsame Struktur zugewiesen werden:

(40)   a.   V/1 $_{\text{FOC}}[\text{Vfin...DR}_{\text{new}}\text{...}]$

        b.   V/2 $_{\text{TOP=BGR}}[\text{DR}_{\text{giv/acc}}]$ $_{\text{FOC=COMMENT}}[\text{Vfin...}]$

Der Wandel in der Verbsyntax vom Ahd. zum Nhd. hin besteht demnach in der Generalisierung des V/2-Musters über den Bereich von Sätzen mit einem expliziten referentiellen Satztopik hinaus sowie in der damit einhergehenden Marginalisierung von V/1-Strukturen auf textinitiale Sätze in metrischer Dichtung und umgangssprachlicher Verwendung.

4   Gegenprobe

Die Zwischenbilanz aus der empirischen Untersuchung der Zusammenhänge zwischen Verbstellung und pragmatischem Status von Diskursreferenten im ahd. Tatian, die eine systematische Präferenz der V/1- und V/2-Verbstellung in unterschiedlichen Kontexten ergab, soll an dieser Stelle einer Gegenprobe unterzogen werden. Zum einen sollen Fälle diskutiert werden, die vorerwähnte und daher allgemein topikfähige Konstituenten enthalten, dennoch aber eine V/1- bzw. *thô*+V/2-Struktur aufweisen. Zum anderen sollen regelmäßige Vorkommen von V/2-Stellungen im Ahd. erörtert werden, in denen die einzige Konstituente vor dem finiten Verb nicht vorerwähnt ist und damit nicht zwingend als Topik im Sinne des aboutness-Konzepts einzuordnen ist. Dazu gehören direkte Fragesätze mit satzinitialem Fragewort und Sätze mit einer nicht-vorerwähnten, daher fokalen Konstituente vor dem finiten Verb.

4.1 Weitere Verwendungsbereiche von V/1 im ahd. Tatian

Auffällig ist, dass die V/1-Stellung im ahd. Tatian über die regelmäßige Verwendung in der Funktion der Ersteinführung von Diskursreferenten hinaus in zahlreichen weiteren Kontexten vorkommt. Interessanterweise enthalten diese Gruppen von Sätzen bereits vorerwähnte Diskursreferenten, welche jedoch nicht in der Position vor dem finiten Verb, sondern rechts davon erscheinen. Dieser Befund scheint das von uns gewonnene Fazit über die Verteilung der Verbstellung in Abhängigkeit von der pragmatischen Funktion der Sätze zunächst zu widerlegen und soll deshalb hier genauer untersucht werden.

Starke Verbreitung findet das V/1-Muster in Sätzen mit Fortbewegungsverben. Ein Teil von ihnen dient in der Tat zur Ersteinführung von Diskursreferenten, vgl. (41):

(41)   **uenit** <u>mulier</u> / de samaria haurire aquam;
         quam tho <u>uuib</u>_{new} / fon samariu sceffen uuazzar (ahd. T 130, 30 f.)
         ‚Kam da eine Frau aus Samaria, um Wasser zu schöpfen'

Die textpragmatische Funktion dieser Äußerung stimmt demnach weitgehend mit der von Existentialkonstruktionen und Präsentationssätzen überein. Das trifft jedoch nicht auf alle Sätze mit Fortbewegungsverben zu. Die V/1-Stellung kommt nämlich auch dann vor, wenn der Diskursreferent in einem Satz mit einem Fortbewegungsverb bereits vorerwähnt ist, vgl. (42)–(43):

(42)   **E**t regressus **est** <u>ihesus</u>
         inti **uuidarfuor** tho <u>ther   heilant</u>_{giv} (ahd. T 53, 14)
         ‚Und der Heiland kehrte da zurück'

(43)   & reuersus **est** <u>centurio</u> in domum suam
         uuarb tho <u>ther centenari</u>_{giv} in sin hús (ahd. T 84, 8)
         ‚Ging da der Zenturio in sein Haus zurück.'

Ferner findet sich die V/1-Stellung im ahd. Tatian regelmäßig in Sätzen, die den Beginn eines neuen Erzählabschnitts signalisieren. Besonders häufig sind darunter die Fälle mit der Floskel ahd. *uuard tho* ‚es geschah, es trug sich zu, etc.' für lat. *factum est.* Wie (44) zeigt, tendiert das Ahd. auch dann zu einer satzinitialen Stellung des finiten Verbs, wenn das unpersönliche Verb *uuard* gemeinsam mit einem Partizip Perfekt auftritt und damit auch eine genaue Nachbildung des Originals möglich gewesen wäre:

(44)   Factum **est** autem in allio sabbato
         uuas thó giuuortan in anderemo sambaztag (ahd. T 106, 6)
         ‚Es war da geworden an einem anderen Sabbatstag'

Auch außerhalb dieser Anfangsfloskel begegnet die V/1-Stellung in Kontexten, in denen ein persönlicher Ausdruck den Beginn eines neuen Erzählabschnitts bezeichnet. Oft entspricht das dem Übergang zu einer neuen Bibelstelle und ist anhand der Konkordanzangaben leicht zu verfolgen. Zusätzlich ist der Beginn neuer Textabschnitte im Original graphisch durch Initialen in Capitalis gekennzeichnet. Obwohl in solchen Fällen die Anfangsstellung des Verbs bereits im Lateinischen dominiert, vgl. (45), lassen sich auch Differenzbelege ermitteln, bei denen am Beginn eines neuen Textabschnitts eine V/1-Abfolge im Ahd. hergestellt wird, vgl. (45)–(47):

(45)   **Rogauit** autem illum quidam / phariseus
         bat inan$_{giv}$ sum / phariseus$_{new}$ (T 126, 1–2)
         ‚Es bat ihn ein Pharisäer'

(46)   Non de omnibus <u>uobis</u> **dico**
         ni **quad** <u>ih</u>$_{giv}$ fon <u>íu</u>$_{giv}$ allen (T 271, 21)
         ‚Nicht sprach ich von euch allen'

(47)   **Acceperunt** autem corpus ihesu
        Intfiengung sie_{giv} tho thes heilantes lichamon (T 321, 29)
        ‚Empfingen sie da des Heilands Leichnam'

Allgemein erstreckt sich die V/1-Stellung auch auf Fälle, die den Beginn neuer
Situationen innerhalb desselben Textabschnitts bezeichnen, so bei inchoativer
Lesart wie in (48):

(48)   Phariseus autem **coepit** intra se / reputans dicere
        bigonda ther phariseus_{giv} innan imo / ahtonti queden (ahd. T 126, 5 f.)
        ‚Es fing der Pharisäer an, bei sich zu sprechen'

Damit lässt sich feststellen, dass die Position des finiten Verbs im Ahd. nicht nur
Leistungen im Bereich der informationsstrukturellen Gliederung der Äußerung
übernimmt, sondern global als textstrukturierendes Signal eingesetzt wird, das
einen Situationswechsel im Diskurs markiert.

    Mit der Diskurssemantik des Situationswechsels sind weitere regelmäßige
Fälle von V/1 im Tatian verbunden, die den Eintritt eines neuen physischen,
psychischen oder kognitiven Zustands an bekannten Diskursreferenten
bezeichnen:

(49)   factus **est** timor super omnes uicinos eorum
        uuard thô forhta ubar alle Iro nahiston_{giv} (ahd. T 31, 2)
        ‚Furcht überkam da alle ihre Verwandten'

(50)   & **sensit** corpore / quod sanata ess& a plaga
        furstuont siu_{giv} thó in ira lihhamen / thaz siu heil uuas fon theru suhti
        (ahd. T 95, 14 f.)
        ‚Da merkte sie an ihrem Körper, dass sie von der Krankheit geheilt
        worden war.

(51)   **Cognouit** ergo <u>pater</u>
         furstuont tho <u>ther fater</u>_{giv} (ahd. T 91, 2)
         ‚Da verstand der Vater'

Schließlich ist bei den V/1-Verwendungen im ahd. Tatian eine große Gruppe an Belegen, darunter auch Differenzbelegen, zu erwähnen, die verba dicendi enthalten und einen Sprecherwechsel im Dialog ankündigen. Auch hier folgen vorerwähnte Diskursreferenten den finiten Verb nach:

(52)   & respondens <u>angelus</u> / dixit ei.
         antlingota tho <u>ther engil</u>_{giv} / quad Iru. (ahd. T 28, 26 f.)
         ‚Da antwortete der Engel [und] sagte zu ihr'

(53)   & respondens <u>mater eius</u> & dixit
         antlingota thô <u>sîn muoter</u>_{giv} Inti quad (ahd. T 30, 24)
         ‚Da antwortete seine Mutter und sagte'

Die Verberststellung in diesen Fällen kann damit erklärt werden, dass eine Redeeinleitung an sich mit dem Wechsel in der generellen Beschaffenheit der Erzählsituation verbunden ist und diese Kontexte dem oben beschriebenen Situationswechsel mit V/1-Stellung funktional nahe kommen.

Es ist offenkundig, dass das V/1-Muster im Ahd. eine ganze Reihe von Verwendungen besitzt, die in den Bereich der Textorganisation und Diskursstrukturierung führen. Die Funktion der Ersteinführung von Diskursreferenten, die oben beobachtet wurde, ist nur eine davon. Mit dieser teilen die im Folgenden aufgeführten weiteren Gruppen von V/1-Belegen auch die funktionsidentische Doppelung durch *thô*+V/2:

(54)   **adducunt** eum / ad pharisęos
         tho **leittun** sie thén / zi then pharisein (ahd. T 221, 15)
         ‚Da führten sie diesen zu den Pharisäern'

(55) & repl&i **sunt** omnes / in sinagoga ira
thó **uuvrdun** sie gifullte alle / in theru samanungu gibuluhti
(ahd. T 115, 7)
‚Da wurden sie alle in der Versammlung von Zorn erfüllt'

(56) **Dixit** autem maria
thô **quad** maria (ahd. T 128, 18)
‚Da sagte Maria'

Die Sichtung des Materials legt nahe, dass bei den verschiedenen Schreibern eine unterschiedlich ausgeprägte Präferenz für *thô*+V/2 vs. V/1 in den Kontexttypen vorliegt, die einen Situationswechsel signalisieren. Eine erste Quantifizierung liegt für den Bereich der Redeeinleitung vor. Diese zeigt eine 100%ige Präferenz für *thô*+V/2 in der Redeeinleitung bei Schreiber ε; der Vergleich der Vorkommen von V/1 vs. *thô*+V/2 in derselben Textmenge bei anderen Schreibern ergab ein Verhältnis 16:3 bei α, 3:9 bei β und 1:12 bei ζ. In Bezug auf diese Fragestellung sind jedoch weitere Nachforschungen erforderlich.

4.2 V/2 in Sätzen mit präverbaler Neuinformation

Die Herausbildung der V/2-Stellung im frühesten Deutsch hängt nach unseren Beobachtungen primär mit dem Bereich der anaphorischen Referenz zusammen, wobei das finite Verb ein Satztopik topologisch vom inhaltlichen Kernstück der Aussage, der Domäne des Neuinformationsfokus trennt. Dieser Befund soll mit anderen Verwendungen der V/2-Stellung im Ahd. verglichen werden, bei denen die Konstituente vor dem finiten Verb nicht topikal ist.

Zum einen handelt sich dabei um direkte Ergänzungsfragen, die gemäß einer detaillierten empirischen Eigenuntersuchung (Petrova & Solf in Vorbereitung) bereits im ältesten Deutsch über eine sehr rigide V/2-Syntax verfügen. Diese zeigt sich darin, dass das Grundmuster bei direkten

Ergänzungsfragen, nämlich die Abfolge Fragewort-V/2-Rest, im Tatian in 202 der insgesamt 230 ermittelten Belege vorliegt, und in 125 dieser Fälle in Differenzbelegen gegen das Original erzeugt wird, vgl. (57) :

(57)   unde hoc **sciam**   wh>DO>Vfin
         uanan **uueiz** ih thaz (ahd T 27, 10)   wh>Vfin>SU>DO
         ‚Woher weiß ich das?'

Entscheidungsfragen, die statt der V/2-Struktur mit satzinitialem Fragewort ein davon abweichendes Muster aufweisen, sind bis auf 4 Differenzbelege von insgesamt 28 Vorkommen stets als Nachbildungen des Originals zu erklären und darüber hinaus vor allem auf einen einzigen Schreiber, nämlich γ, beschränkt.

Daraus folgt, dass sich im Bereich der direkten Ergänzungsfragen offenbar seit frühester Zeit eine feste V/2-Syntax ausgebildet hat. Das Problem dieses Fazits für die vorliegende Arbeit besteht darin, dass die präverbale Konstituente in direkten Ergänzungsfragen, d. h. das Fragewort selbst, nicht mit pragmatischen und informationellen Merkmalen von Topikalität in Verbindung gebracht werden kann. Wir finden m. a. W. in gewissen Bereichen des Ahd. Evidenz für eine fest ausgebildete V/2-Syntax, die mit den empirischen Beobachtungen aus der informationsstrukturellen Analyse anderer früher Vorkommen von V/2 in Deklarativsätzen nicht vereinbar ist.

Probleme scheinen ferner solche V/2-Strukturen in Deklarativsätzen zu bereiten, in denen die präverbale Konstituente neue Information trägt und daher nicht als topikal, sondern als fokal zu deuten ist. Dies ist etwa in dem folgendem Differenzbeleg der Fall, wo wir links vom finiten Verb die Domäne des engen XP-Fokus finden, rechts davon die topikale Konstituente ahd. *her* ‚er':

(58)   mortuus **est**
         toot **ist** <u>her</u>$_{giv}$ (ahd. T 149, 14)
         ‚Tot ist er'

Im Vergleich wird allerdings deutlich, dass die Domäne des Neuinformationsfokus im Ahd. in einer verhältnismäßig größeren Anzahl an Differenzbelegen rechts vom finiten Verb realisiert wird, vgl. (63)–(66):

(59)   quinque enim uiros **habuisti**
         <u>thu</u>$_{giv}$ **habetos** <u>finf gomman</u>$_{new}$ (ahd T 132, 6)
         ‚Du hattest fünf Männer'

(60)   alii autem nequaquam. / sed similis **est** eius
         andere quadun nist / úzouh <u>hér</u>$_{giv}$ **ist** imo gilih (ahd. T 221, 5–6)
         ‚Andere sagten: „Er ist es nicht, aber er ist ihm gleich."'

(61)   <u>lutum</u> **fecit**
         <u>her</u>$_{giv}$ **t&a** <u>leimon</u>$_{new}$ (ahd. T 221, 10)
         ‚Er rührte Schlamm ein'

(62)   <u>demonium</u> **hab&**
         <u>er</u>$_{giv}$ **hab&** <u>diuual</u>$_{new}$ (ahd. T 226, 18)
         ‚Er ist vom Teufel besessen'

Schließlich ist ein kleiner Bereich von V/2-Strukturen in Differenzbelegen des ahd. Tatians anzuführen, bei dem sich vor dem finiten Verb ein kataphorisches Element befindet, das auf einen postverbalen weiten Fokus verweist (Fokusvorwegnahme):

(63)   & hoc uobis signum, Inueni&is / Infantem pannis Inuolutum.' / &
         positum in presepio,
         thaz **sî** îu zi zeichane. thaz ir find& / kind mit tuohon biuuvntanaz.' / Inti
         gilegitaz in crippa; (ahd. T 36, 8 ff.)
         ‚Das sei euch zum Zeichen, dass ihr ein Kind finden werdet, das in
         Tücher gewickelt und in eine Krippe gelegt ist'

Die Katapher ahd. *thaz* ‚das' ist Gegenstand einer weiterführenden Erläuterung. Vor diesem Hintergrund kann argumentiert werden, dass das kataphorische Element informationsstrukturell betrachtet Merkmale eines aboutness-Topiks aufweist, sofern es eine Entität etabliert, über die eine Aussage getroffen wird.

5   Ausblick und weitere Untersuchungen

Das Fazit der empirischen Untersuchung über die Korrelation von Verbstellung und pragmatischem Status von Diskursreferenten lautete, dass die V/2-Stellung ein konstitutives Merkmal von Sätzen mit bekannten Diskursreferenten darstellt und in funktionaler Opposition zu V/1-Sätzen steht, in denen Diskursreferenten erstmals eingeführt werden. Die Gegenprobe zur Informationsstruktur bei davon abweichenden V/1- und V/2-Vorkommen in Differenzbelegen des ahd. Tatians erfordert an dieser Stelle einige Präzisierungen.

Zunächst ist für die Fälle von V/1 davon auszugehen, dass dieses Verbstellungsmuster eine breite Verwendung im Ahd. fand und in einer größeren Anzahl an Kontexten bevorzugt erscheint, die ganz allgemein einen Situationswechsel im Diskurs signalisieren. Das V/2-Muster dagegen ist auch in Differenzbelegen nachzuweisen, bei denen das finite Verb keine referentielle bzw. keine topikale Vorfeldkonstituente vom Neuinformationsfokus der Aussage trennt. Dazu zählen die Frame-Adverbiale *thô, thar* sowie die kataphorischen Elemente, die auf nachfolgende Neuinformation verweisen. Sie zeugen davon, dass bereits in ahd. Zeit ein Wandel im Gang ist, der sich in der Ausbreitung des V/2-Musters über die Domäne der anaphorischen Referenz hinaus äußert. Als ein Zeichen der fortschreitenden Ausbreitung von V/2 sind auch die festgestellten Fälle von engen XP-Foki links vor dem finiten Verb zu

deuten, die allerdings bei den Differenzbelegen zahlenmäßig deutlich hinter Fällen von postverbalem Fokus zurückstehen.

Bedeutsamer für das Verständnis der Generalisierung von V/2 ist die Evidenz bei den selbständigen Fragesätzen mit Fragewort. Für ihre Erklärung kommen aus unserer jetzigen Sicht zwei Möglichkeiten in Betracht. Die erste setzt eine Modifizierung der These von der Funktion des finiten Verbs als Mittel der Trennung der informationsstrukturellen Domänen von Topik und Fokus in Satz voraus. Eine mögliche Anpassung dieser These zielt auf die Generalisierung der Funktionsweise des finiten Verbs im Ahd. Demnach bewirkt das finite Verb die Abgrenzung einer pragmatisch konsistenten Vorfelddomäne, die entweder topikal oder fokal, jedoch nicht pragmatisch gemischt sein kann. Eine ähnliche informationsstrukturelle Differenzierungsfunktion von Adverbien im Mittelfeld, die eine Domänenkonsistenz bei der Serialisierung von Argumenten schaffen, beobachtet Musan (2002) im heutigen Deutsch. Wir können in Analogie dazu die Stellung des finiten Verbs im Ahd. als ein Mittel erklären, informationsstrukturelle Domänen nach Möglichkeit als zusammenhängende Einheiten voneinander zu trennen.

Die zweite Möglichkeit, die Generalisierung von V/2 mit der Evidenz im Bereich der direkten Fragesätze in Einklang zu bringen, geht mit der Vermutung einher, dass die Entwicklung der V/2-Stellung bei den direkten Fragesätzen und bei den Deklarativsätzen das Ergebnis zweier voneinander unabhängiger Prozesse darstellt: eines syntaktisch bedingten Prozesses bei den Fragesätzen und eines informationsstrukturell bedingten Prozesses bei den Deklarativa mit V/2-Abfolge in kategorischer Lesart. Der unterschiedliche pragmatische Status der präverbalen Elemente in diesen Bereichen fester V/2-Setzung führt zu einer Umdeutung der Vorfeldposition bei den Deklarativa. Diese wird von einer

informationsstrukturell vorbestimmten zu einer informationsstrukturell indifferenten Domäne, wobei nunmehr allein die Anzahl, jedoch nicht die Pragmatik der sie besetzenden Konstituenten von Belang ist.

Wie die Daten zeigen, ist der Generalisierungsprozess von V/2 in dem untersuchten Zeitraum bereits in Gang. Das erklärt die informationelle Vielfalt der präverbalen Elemente bei V/2-Abfolgen in Differenzbelegen des ahd. Tatians. Gleichzeitig sind jedoch klare quantitative Präferenzen zugunsten referentieller Topikelemente im Vorfeld von V/2-Strukturen zu erkennen. Dies deutet sowohl auf den funktionalen Ursprungsbereich von V/2 bei den Deklarativa mit kategorischer Lesart hin, als auch auf den Wandelprozess, der sich in der kontinuierlichen Ausbreitung des V/2-Musters auf weitere Kontextbereiche äußert.

6   Anhang

exurgens autem ioseph asomno
fecit sicut precepit ei angelus dni
& accepit coniugem suam
& non cognoscebat eam donec peperit
filium suum primogenitum .
& uocauit nomen eius ihm ,
Factum est autem Indiebus illis.
exiit edictum acesare augusto .
utdescriberetur uniuersus orbis ,
hec descriptio prima facta est
preside syriae cyrino .
& ibant omnes ut profiterentur
singuli insuam ciuitatem ,
Ascendit autem & ioseph agalilea
deciuitate nazareth .
Iniudaeam ciuitatem dauid
quae uocatur bethleem ,
eoquod esset dedomo
& familia dauid . ut profiteretur
cum maria desponsata sibi
uxore pregnante .
factum est autem cu essent ibi .
Impleti sunt dies ut pareret .
& peperit filium suum primogenitu ,
& pannis eum inuoluit
& reclinauit eum inpresepio
quia non erat eis locus
Indiuersorio ,
Et pastores erant Inregione eadem .
uigilantes &custodientes uigilias noctis
supragrege suum.
& ecce angelus dni stetit iuxta illos .

dô stantanti thô ioseph fon slafe
tëtâ sô imo gibôt truhtines engil
inti Infieng sina gimahhun
inti niuuard Iru uuar unzan siu gibar
Irasun êrist boranon
inti gihiez sinan namon heilant ,
uuard thô grân Inthên tagun
fram quâ gibot fon demo Al uualten keisure
thaz gibriebit uurdi Al thesa umbi uuerft ,
thaz giscrib Iz êristen uuard grân
Insyriu fon demo grauen cyrine .
Inti fuorun alle . thaz bi iâhin
thionost Insinero burgi ,
fuor thô ioseph fon galileu
fon thero burgi thiu hiez nazareth
In Iudeno lante Inti In dauuider burg
thiu uuar ginemnit bethleem .
bithiu uuanta her uuar fon huse
Inti fon hiuuiske dauides . thaz her gilahi
saman mit mariun Imo gimahaltero
gimahhun sô raffânteru .
thô sie thar uuarun .
uurdun taga gifulte thaz siu bari
inti gibar Irasun êrist boranon .
Inti biuuant Inan mit tuochan
Inti gilegita Inan Incrippea
bithiu uuanta In niuuar andero stat
Inthemo gast huse .
uuarun thô hirta Inthero lante skeffi .
uuahhante Inti bihaltante naht uuahta
ubar ero euuit ,
quam thara gotes engil Inti gistuont nâh In

Reference

Textquellen

[T]   Die lateinisch-althochdeutsche Tatianbilingue Stiftsbibliothek St. Gallen
        Cod. 56. Hg. v. A. Masser. Göttingen: Vandenhoeck & Ruprecht, 1994.
        (= Studien zum Althochdeutschen, Bd. 25)

Wissenschaftliche Literatur

Bean, Marian C. 1983. The Development of Word Order Patterns in Old
        English. London and Canberra: Croom Helm. (=Croom Helm Linguistics
        Series)

Behaghel, Otto. 1932. Deutsche Syntax. Eine geschichtliche Darstellung. Band
        IV. Wortstellung. Periodenbau. Heidelberg: Winter. (= Germanische
        Bibliothek. I. Sammlung germanischer Elementar- und Handbücher. 1.
        Reihe: Grammatiken. 10. Bd.)

Braune, Wilhelm. 1894. Zur Lehre von der deutschen Wortstellung. In:
        Forschungen zur deutschen Philologie. Festgabe für Rudolf Hildebrand
        zum 13. März 1894. Leipzig: Veit & Co.

Dik, Simon C. 1997. The Theory of Functional Grammar. Part I: The Structure
        of the Clause. Second, revised Edition. Edited by K. Hengeveld. Berlin /
        New York: Mouton de Gruyter. (= Functional Grammar Series 20)

Dittmer, Arne und Ernst Dittmer. 1998. Studien zur Wortstellung.
        Satzgliedstellung in der althochdeutschen Tatianübersetzung. Göttingen:
        Vandenhoeck & Ruprecht. (= Studien zum Althochdeutschen, Bd. 34)

Donhauser, Karin und Roland Hinterhölzl. 2003. Die Rolle der
        Informationsstruktur bei der Herausbildung von Wortstellungsregularitäten
        im Germanischen. In: C. Féry et al. (Hgg.): Informationsstruktur. Die
        sprachlichen Mittel der Gliederung von Äußerung, Satz und Text.
        Finanzierungsantrag 2003–2007. Potsdam und Berlin: Universität Potsdam,
        173–193.

Drubig Bernhard. 1992. Zur Frage der grammatischen Repräsentation thetischer
        und kategorischer Sätze. In: J. Jacobs (Hg.): Informationsstruktur und
        Grammatik. Linguistische Berichte. Sonderheft 4, 142–195.

Fourquet, Jean. 1974. Genetische Betrachtungen über den deutschen Satzbau. In: Besch, Werner et al. (Hgg.) Studien zur deutschen Literatur und Sprache des Mittelalters. Festschrift für Hugo Moser zum 65. Geburtstag. Berlin: E. Schmidt, 314–323.

Frey, Werner. 2000. Über die syntaktische Position des Satztopiks im Deutschen. In: ZAS Papers in Linguistics 20, 137–172.

Karttunen, Laurie. 1976. Discourse referents. In: J. McCawley (Hg.): Syntax and Semantics, Vol. 7. New York: Academic Press, 363–385.

Kiparsky, Paul. 1995. Indoeuropean Origins of Germanic Syntax. In: A. Battye and I. Roberts (eds.): Clause Structure and Language Change. Oxford: Oxford University Press, 140–169.

Klein, Wolfgang und Christiane von Stutterheim. 1992. Textstruktur und referentielle Bedeutung. In: Zeitschrift f. Literaturwissenschaft und Linguistik 86, 67–92.

Lambrecht, Knud. 1994. Information Structure. Topic, Focus, and the Mental Representations of Discourse Referents. Cambridge: Cambridge University Press. (= Cambridge Studies in Linguistics, 71)

Lenerz, Jürgen. 1984. Syntaktischer Wandel und Grammatiktheorie. Eine Untersuchung an Beispielen aus der Sprachgeschichet des Deutschen. Tübingen: Niemeyer.

Lühr, Rosemarie (im Druck): Der Einfluß der klassischen Sprachen auf die germanische Grammatik. In: G. Meiser, Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft in Halle/S., 17. –23. September 2000.

Masser, Achim. 1997a. Syntaxprobleme im althochdeutschen Tatian. In: Y. Desportes (Hg.), Semantik der syntaktischen Beziehungen. Akten des Pariser Kolloquiums zur Erforschung des Althochdeutschen 1994. Heidelberg: Winter, 123–140.

Masser, Achim. 1997b. Wege zu gesprochenem Althochdeutsch. In: E. Glaser und M. Schlaefer (Hgg.): Grammatica Ianua Artium. Festschrift für Rolf Bergmann zum 60. Geburtstag. Heidelberg: Winter, 49–70.

Musan, Renate. 2002. Informationsstrukturelle Dimensionen im Deutschen. Zur Variation der Wortstellung im Mittelfeld. In: ZGL 30, 2002, 198–221.

Petrova, Svetlana und Michael Solf (in Vorbereitung): Syntaktischer Wandel und Satzmodus. Beobachtungen zur Wortstellung in direkten Fragesätzen des Althochdeutschen.

Prince, Ellen F. 1981. Towards a Taxonomy of Given-New Information. In: P. Cole (Hg.), Radical Pragmatics. New York: Academic Press, 223–255.

Ruhfus, Wilhelm. 1897. Die Stellung des Verbums im althochdeutschen Tatian. Inaugural-Dissertation zur Erlangung der Doctorwürde. Dortmund: Selbstverlag.

Schlachter, Eva. 2004. Satzstruktur im Althochdeutschen. Eine Skizze zur Position des Verbs im Isidor-Traktat des 8. Jahrhunderts. In: K. Pittner et al. (Hgg.), Beiträge zu Sprache und Sprachen 4. Vorträge der Bochumer Linguistik-Tage. München: Lincom, 179–188.

Wackernagel, Jacob. 1891/92. Über ein Gesetz der indogermanischen Wortstellung. In: Indogermanische Forschungen 1, 333–346.

Roland Hinterhölzl
Humboldt-Universität zu Berlin
Institut für deutsche Sprache und Linguistik
Sitz: Schützenstr. 21
Unter den Linden 6
10099 Berlin
roland.hinterhoelzl@rz.hu-berlin.de

Svetlana Petrova and Michael Solf
Humboldt-Universität zu Berlin
SFB 632 „Informationsstruktur"
Sitz: Mohrenstr. 40-41
Unter den Linden 6
10099 Berlin
Germany
www2.hu-berlin.de/sprachgeschichte/forschung/informationsstruktur

Focus accent, word length and position as cues to L1 and L2 word recognition [*]

Anke Sennema, Ruben van de Vijver,
Susanne E. Carroll, Anne Zimmer-Stahl

Universität Potsdam

The present study examines native and nonnative perceptual processing of semantic information conveyed by prosodic prominence. Five groups of German learners of English each listened to one of 5 experimental conditions. Three conditions differed in place of focus accent in the sentence and two conditions were with spliced stimuli. The experiment condition was presented first in the learners' L1 (German) and then in a similar set in the L2 (English). The effect of the accent condition and of the length and position of the target in the sentence was evaluated in a probe recognition task. In both the L1 and L2 tasks there was no significant effect in any of the five focus conditions. Target position and target word length had an effect in the L1 task. Word length did not affect accuracy rates in the L2 task. For probe recognition in the L2, word length and the position of the target interacted with the focus condition.

Keywords: bilingual word processing, prosodic prominence

1 Introduction

Focus, as expressed through pitch, has been identified as an important aspect of language comprehension, and a number of factors influence the assignment of focus during sentence comprehension.

It has been shown that rapid and effective processing of the accent placement in an utterance contributes to efficient comprehension of meaning (Cutler & Fodor,

1979) and that there is a robust advantage for words with predicted accent (Cutler, 1976). Listeners exploit cues in the prosodic contour preceding an accent to locate possible accent points. Since accent falls on semantically crucial, i.e. focused words, listeners actively search for focus to facilitate comprehension. Drawing the listeners' attention to certain parts of the utterance through means of prosodic prominence might lead to more detailed processing of the signal and this might lead to faster word recognition (Cutler *et al.*, 1997).

This evidence from studies of native listening inspired investigations addressing nonnative processing strategies of accentual sentence structures. A consistent and similar interaction of accent and focus was reported for native processing of English and Dutch, although this effect did not show in nonnative listening (Akker & Cutler, 2003). The recognition memory of English sentences was tested in which prosody cued meaning contrasts, and memory performance based on prosodic information was shown to be generally poor (Pennington & Ellis, 2000). After participants' attention was explicitly directed to intonation, the performance improved only on sentences with contrastive focus pairs.

From this it is not yet clear, to what extent the prosodic realization of focus play a role in L2 processing. Therefore, in a comparative analysis we investigated the role of focal accent, word length and the position of the word in the sentence on word recollection in native and nonnative recognition tasks in German L2 learners of English. If it is the case that the realization of focus through pitch enhances processing and recollection than such effects occur in short and long words and independent of the position in the sentence of the target word.

2   Study

We aimed at examining the relationship between effects of focus as realized by accent for native and nonnative listening. Furthermore, we have addressed the influence of sentence position and word length on word recognition in L2.

For L1 it has been established that accented material is processed more efficiently (Cutler, 1976). It has been shown, that L2 listeners attend more readily to the beginning and end of the sentence and that learners show a preference for the sentence beginning (Barcroft & VanPatten, 1997), (Klein, 1992). We will investigate the extent to which prosodic prominence interacts with the Sentence Location Principle of VanPatten. Our hypothesis is, in contrast to VanPatten, that learners will be most sensitive to words which are prosodically marked for focus regardless of its location in the sentence. We expect, therefore, that German L2 learners of English will recognize target words better when these occur at sentence initial or final position.

It has been argued, that stressed syllables are articulated in a clearer fashion (Cutler & Norris, 1988), and that hence it would be easier to represent its phonetics and phonological characteristics. In the case of words of more than one syllable, the listener might be able to accurately represent the properties of the stressed syllable but not of unstressed syllables, making recognition of the whole word more difficult than recognition of a word consisting of one stressed syllable. All of these issues - the influence of focus as realized by accent, position in the sentence and word length - have been addressed in the present study.

2.1   Speech materials

Target words in each language consisted of names of birds, such as *Gohl* (a German target) or *scaups* (an English target). They were judged by three German learners of English for their familiarity and balanced in the lists of

target words. The items were one-syllabled or more-syllabled. For every target one filler item was constructed. Target items were controlled for word length and sentence position, the fillers were not. The items were embedded in sentences.

To study the effects of different focus positions on word recognition, prosodic variation was triggered by using wh-questions for otherwise identical sentences. A wh-question focuses a specific constituent of the sentence and the answer to that question ought to focus the same constituent (Selkirk, 1992). The following example from the test material shows the relation between pitch accent and focus (pitch accent is marked by capital letters):

(1) a.   "AUKS are being affected by the warming of the northern seas because they prefer cold waters. "

would be the answer to the question

     b.   "Who is being affected by the warming of the northern seas?",

whereas

(2) a.   "Auks are being affected by the warming of the northern seas because they prefer COLD WATERS."

would be the answer to the question

     b.   "Why are auks being affected by the warming of the northern seas?"

Elicited sentences were broad focus sentences (condition B1), narrow focus realised on the target (condition N1, as in example 1a), or narrow focus realised on a constituent other than the target (condition B2, as in example 2a). In the carrier sentence the target words occurred in three different positions, i.e. initial, medial and final.

To evaluate the effect of the cues in the prosodic contour surrounding an accented word, two test conditions with spliced material were constructed. In the

first spliced condition (spliced B1), the target word of a B1 sentence was spliced into the context of a N1 sentence, i.e. an unaccented target word was spliced into the N1 context of focused constituent. In the second condition (spliced B2), the target of a B2 sentence was spliced into a N1 sentence, i.e. an unaccented and unfocused target word was spliced into a sentence with narrow focus on the target. With this material 5 separate experimental conditions with identical set-up were programmed.

2.2   Speakers and Recording procedure

A female speaker of South Eastern British English recorded the English items and a female native speaker of Standard German recorded the items in German.

For each language 24 target items and 24 foils across sentence position and syllable length were recorded in three conditions, i.e. B1, N1 B2. 75 filler sentences were constructed in B1 and B2 sentence conditions, some of which had to be used twice in the experiment to make up numbers. In addition, one token of each target and foil was recorded in isolation. Also, 12 sentences plus 3 target items were recorded for a familiarization phase at the beginning of the experiment. Digital recordings were made in a sound-proof booth, using an Audiotechnica 4033a microphone (audio sampling rate 22.05 kHz).

2.3   Listeners

104 listeners participated in the experiment. They were German students or employees at the University of Potsdam and were at an intermediate to advanced level of English proficiency. The range of age was between 18 and 43 years. All participants had all started learning English after the age of 8 and none had lived in an English speaking country for more than 24 months.

A British English control group of 45 students was tested in the UK. 28 of them were recruited from University College London and 17 from Essex

University, where they were first and second year students of Linguistics.

All participants in both language groups either received points for course requirements or were paid a small sum for their participation. At the time of the experiment no reported normal or corrected hearing and normal or corrected vision.

There were 20 German subjects in each condition N1, B1, B2 and B2 spliced and 24 German subjects in condition B1 spliced. In each condition N1, B2, B1 spliced there were 10 English controls and 13 subjects in condition spliced B2.[1]

2.4   Experimental task and procedure

A closed-set word probe detection task was built for the experiment. The task was explained to the listeners and they were instructed to pay attention to the sentences in order to do well on the word probe detection task. After that they entered a brief training session in which they heard three blocks each of 4 sentences followed by a single word. Listeners were asked to press one key on a computer keyboard when they recognized the word heard in isolation as one having occurred in one of the previous 4 sentences, and another key when the word in isolation did not occur in one of the previous 4 sentences. They were instructed to make their decision as quickly as possible.

In the trial part feedback was given on the correctness of the answers but no feedback was given during the actual test and there was no further communication with the experimenters. Subjects heard the stimuli sentences only once. Two self-timed pauses were programmed within the experiment. Each experiment (German task and English task) took about 30 minutes.

---

[1]   Due to an error only two subjects were tested for condition B1 unspliced.

The experiment had two separate parts: a part with German stimuli and one with English items. Both parts consisted of 48 blocks of four sentences each. There were 24 blocks with 'no' as the correct answer and 24 blocks with 'yes' as the correct answer. The targets were equally distributed in terms of presentation order within the blocks (pos. 1-4) and the blocks were presented in randomized order. To compensate for fatigue effects a second list was created with the blocks in the reversed order of the first list. The sentences and probes were presented at a comfortable listening level via headphones. Their responses and reaction times were recorded.

To obtain an independent measure of English proficiency, the Oxford Placement Test was administered. This is a standardized test (multiple choice) divided into two sections, Listening and Grammar. The percentage scores from the sections were used to obtain a relative measure of English proficiency of the German participants. An attempt was made to stratify our learners by proficiency level on the basis of the score achieved in the Listening part of the Oxford Placement Test and to distribute learners equally across the five focus conditions.

The experimental order was as follows: Oxford Placement Test, German part of the experiment (L1 task), and one week later, the English part of the experiment (L2 task). The English control group was tested on the English part only.

2.5 Results

Timed-out responses (a response latency of more than 2500 ms) and responses with reaction times below 300ms were discarded from the analysis. The percentage of correct probe recognition in each part of the experiment (L1 task and L2 task) was calculated for the five test conditions (Table 1).

Tab. 1: Scores (% correct) per focus condition and part of experiment for
          German subjects and English controls

*N=2

German subjects	**B1**	**N1**	**B2**	**B1 spliced**	**B2 spliced**
L1 task	93,1%	93,7%	92,9%	93,2%	94,2%
L2 task	82,6%	88,9%	85,7%	87,3%	85,3%
Controls	100%*	93,5%	86,3%	89,4%	85,4%

Accuracy scores were subjected to an analysis of variance with language as the within subjects factor (L1 vs L2) and focus as the between subjects factor. There was a significant difference between L1 and L2 [$F(1,99) = 117.21$, $p < .001$], showing that German subjects performed better in their native language L1 than in L2. Whereas there was no main effect of focus, focus tended to interact with language [$F(4,99) = 2.17$, $p = .08$]. Accuracy scores for the experimental task in the L1 and the task in the L2 in the five focus conditions are shown in Figure 1.

Fig. 1: Boxplots of accuracy scores
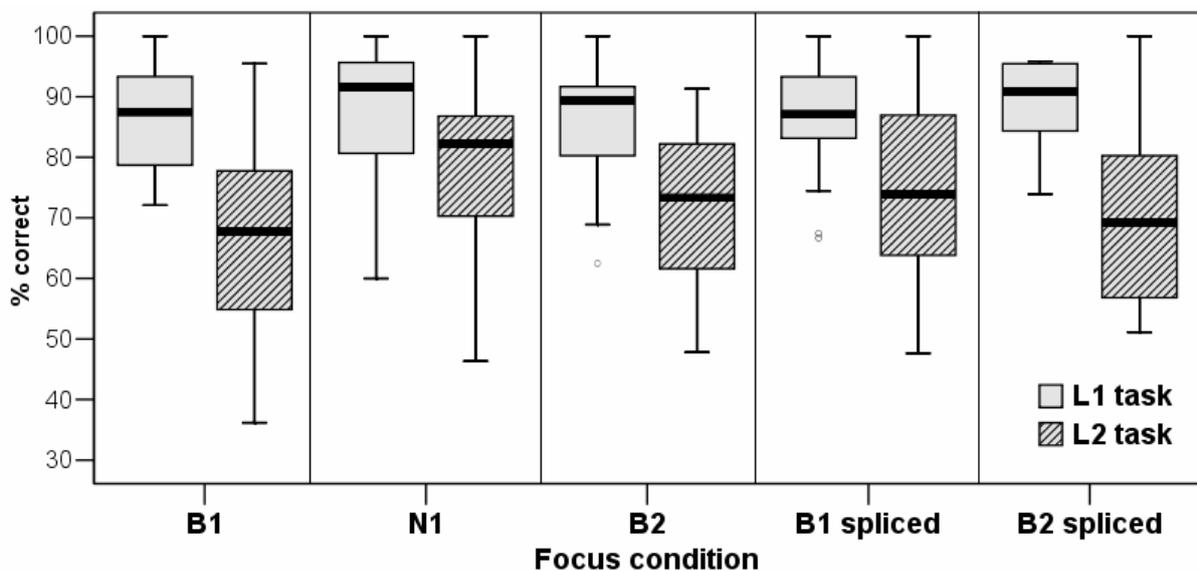
Next, the factors of word length and target position in the sentence were examined for the five conditions in each of the two experimental tasks. Accuracy scores for targets occurring in sentences were subjected to an analysis of variance with target length (mono- or polysyllabic) and target position in the sentence (initial, medial, final) as within subjects factors and focus as between subjects factor. Paired comparisons for target position examined initial and medial position against final position.

In the L1 task with broad focus (B1) there was no effect for target length, but a significant difference between the target positions [$F(2,38) = 6.032$, $p = .05$]. Paired comparisons of the final position against initial and medial position revealed a significant difference between initial and final target position [$F(1,19) = 6.032$, $p < .005$] . In the corresponding L2 task, however, there was no effect for target length and also no effect for target position (see Fig. 2).

Fig. 2: Performance of German subjects in condition B1 unspliced
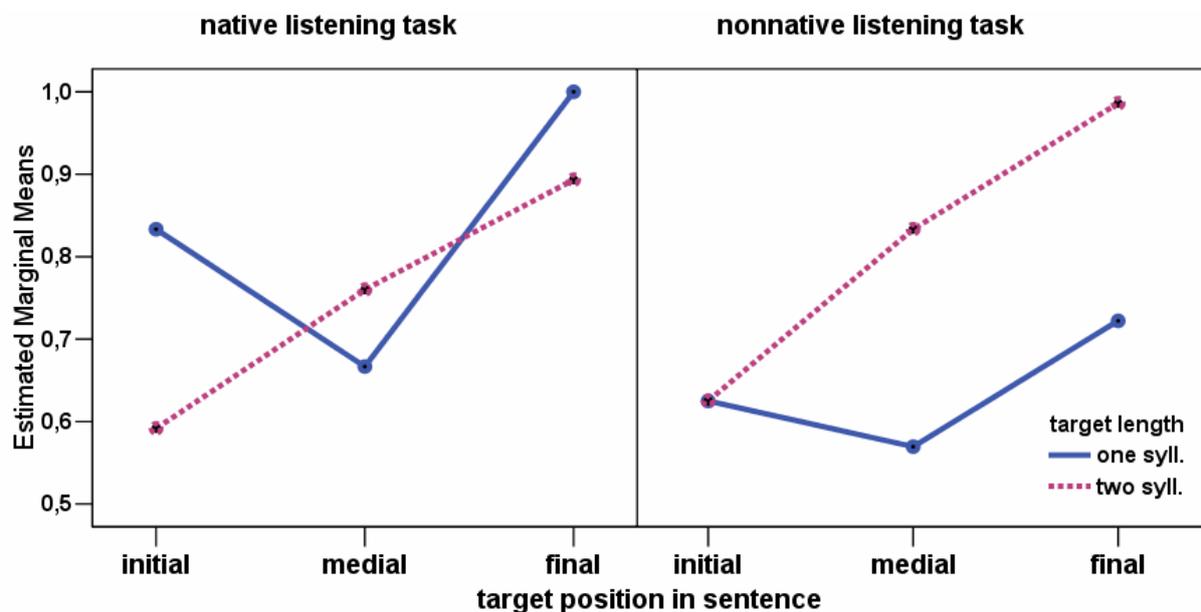
In the condition with an accented target (N1), target length did not have an effect on word recognition in either task. In the native listening task, target

position was significant [$F(2,38) = 5.209$, $p = .01$] and interacted with word length [$F(2,38) = 4.358$, $p < .05$]. Paired comparisons revealed a significant difference between initial and final target position [$F(1,19) = 13.470$, $p < .005$]. There was no effect of target position and no interaction with word length in the L2 task. It appeared that the two patterns of responses in condition N1 are differed regarding the final position in the L1 and L2 task (Fig. 3). Final position does not seem to draw additional attention to targets in second language processing whereas this cues better recognition in the first language.

Fig. 3: Performance of German subjects in condition N1 unspliced
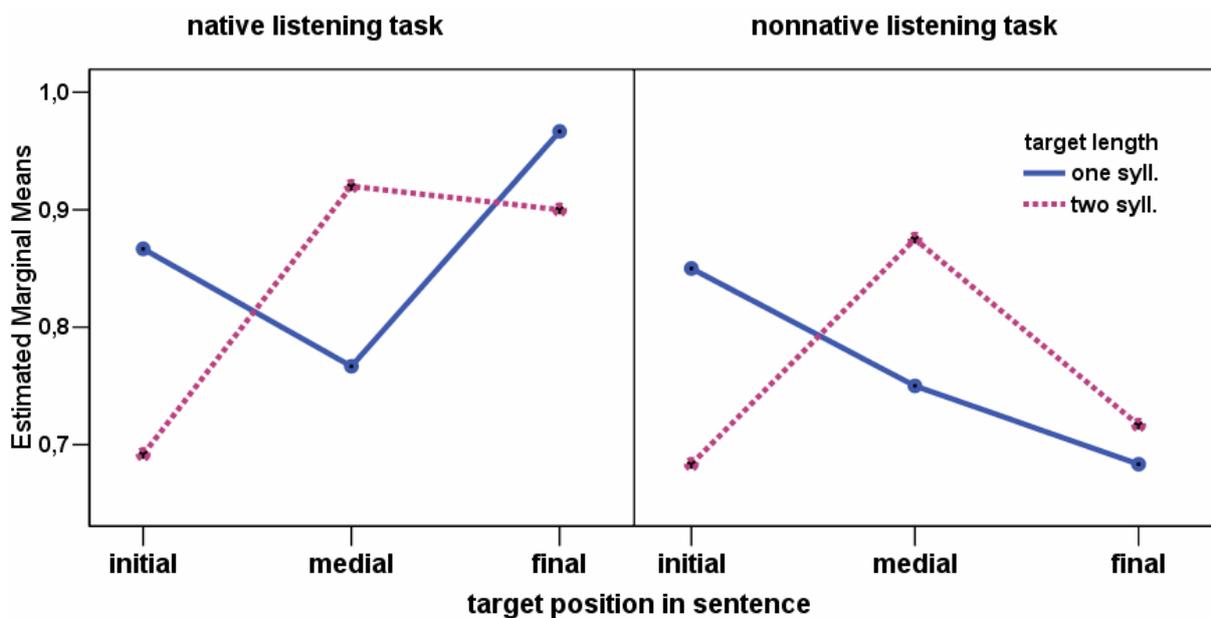
In condition B2 (accent on non-target word in constituent that has the target) there was no effect of word length in the two tasks. In the native listening task, target position was significant [$F(2,38) = 3.261$, $p < .05$] and interacted with word length [$F(2,38) = 4.166$, $p < .05$]. Paired comparisons revealed significant differences between initial and final position of the target [$F(1,19) = 6.265$, $p < .05$] and also between medial and final position [$F(1,19) = 4.388$, $p = .05$]. In

the L2 task the effect of target position was significant $[F(2,38) = 5.028, p < .05]$ and paired comparisons showed a significant difference between the initial and final position $[F(1,19) = 14.241, p < .005]$. There was no interaction of target position with word length in the L2 task.

The response patterns for condition B2 unspliced are shown in Fig. 4.

Fig. 4: Performance of German subjects in condition B2 unspliced
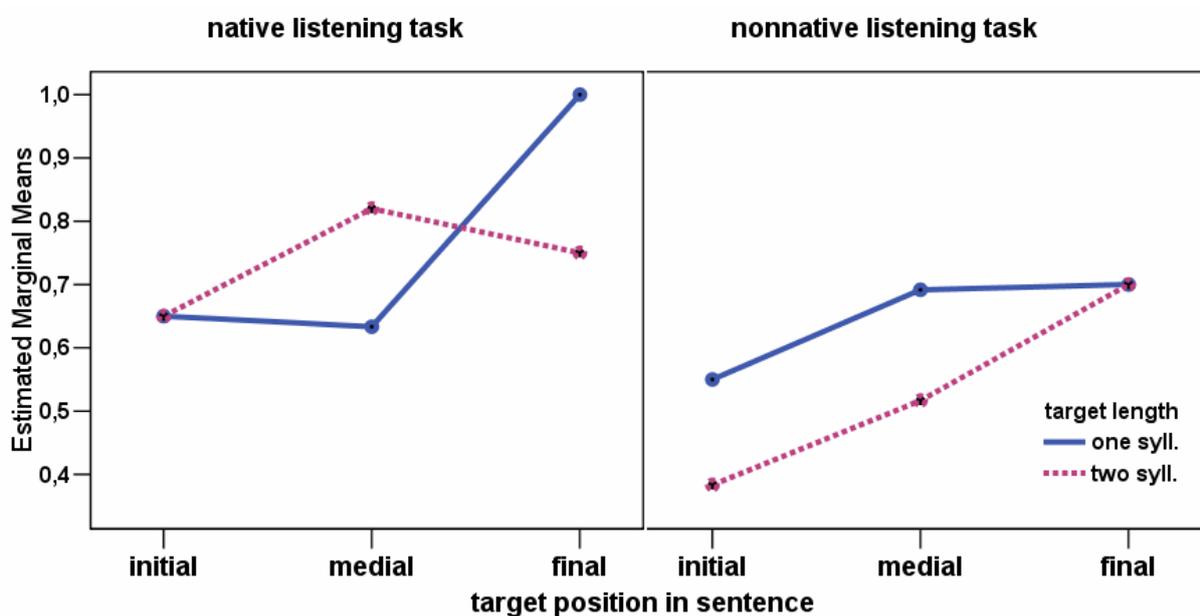
Accuracy scores of the spliced conditions were subjected to the same analyses as described above. In the condition B1 spliced (target word of a B1 broad focus sentence spliced into the context of a N1 narrow focus sentence), there was a significant effect in the L1 task for both target length $[F(1,23) = 26.763, p < .001]$ and for target position $[F(2,46) = 3.253, p < .005]$. There was no interaction between the two factors. In the L2 task of condition B1 spliced, only target position had a significant effect $[F(2,46) = 5.153, p = .01]$ and for this, paired comparisons revealed significant differences between initial and final position $[F(1,23) = 8.013, p < .01]$ and between medial and final position of the

target [$F(1,23) = 10.903$, $p < .005$]. The patterns of responses in condition B1 spliced are similar in the L1 and L2 task, as can be seen in Figure 5.

Fig. 5: Performance of German subjects in condition B1 spliced
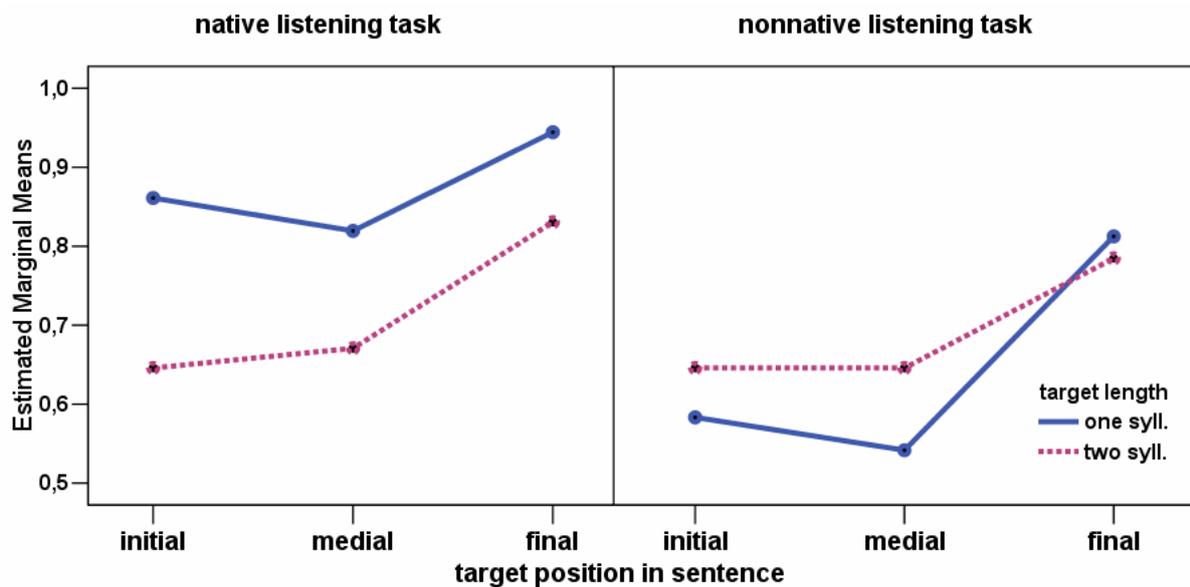
In condition B2 spliced, the target had been spliced out of a B2 sentence (with accent on another word in the constituent that contained the target) into a N1 sentence. In the L1 task there was a significant effect for word length [$F(1,19) = 5.391$, $p < .05$] and word length interacted with target position [$F(2,38) = 6.027$, $p = .005$]. In the L2 task, only target position had a significant effect [$F(2,38) = 3.234$, $p = .05$]. This was due to a significant difference between the medial and final position [$F(1,19) = 6.883$, $p< .05$] (see Fig 6).

Fig. 6: Performance of German subjects in condition B2 spliced
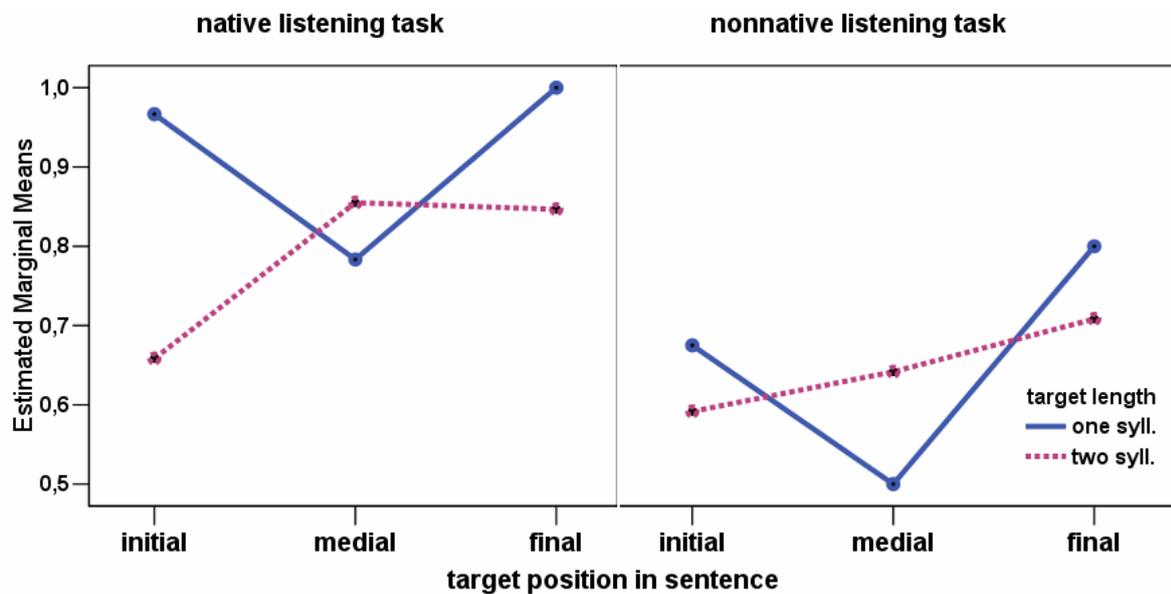
It could be argued from the results that the effect of position found for initial and final position obliterates any possible effect of focus. We therefore examined the recognition accuracy of targets in medial position, thus excluding a superimposed effect of sentence position. Irrespective of word length a stronger effect of focus condition was revealed for targets not occurring at the outer ends of the sentences, where a position effect seems indeed to overshadow focus effect.

Tab. 2: Scores (% correct) per focus condition and part of experiment for targets in medial position

German subjects	B1	N1	B2
L1 task	88,5%	92,2%	88,3%
L2 task	78,9%	90,5%	77,6%
Controls	no values	92,1%	84,2%

3   Discussion

It turns out that accent did not help our subjects process the accented word probes more accurately or more efficiently in L2. There was no significant effect for focus condition. This result confirms the findings of Akker & Cutler (2003) and of Pennington & Ellis (2000). Processing a second language might take up so many resources that accent alone is not a strong enough help. The fact, however, that accent did not help listeners in their L1 either is surprising. This might be due to an effect of target position which is possibly stronger than an effect of accent. This is suggested by the fact that target position had an effect in all accent conditions.

The most salient results were the effects of the position of the target. Over all conditions it seems that position is the strongest cue for the word probe detection task: words in initial or final position of the sentence were remembered better and this confirmed the primacy & recency effect found in other studies (Klein). Our results indicate that in order to effectively commit a representation to memory, the position of the word in the sentence is more important than whether or not a word is accented. It also suggests that effective and efficient processing of accented words (Cutler, 1976) in L1 does not necessarily result in successfully committing the word to memory.

We had hypothesized that learners will be most sensitive to words which are prosodically marked for focus regardless of its location in the sentence and had expected, that German L2 learners of English will recognize target words better when these occur at sentence initial or final position. The effect of sentence position of the target was nevertheless clearly more strong than that of focus accent, which seemed to be merely supportive. Examining the results for target words in sentence medial position only, focus appeared to have indeed a

stronger effect.

The splicing procedure revealed no effects for the cues in the prosodic contour surrounding an accented word and the two test conditions did not lead to conclusive results. In the spliced conditions, target position had a significant effect in L2 and L1.

4   Conclusions

In this study, an experiment was conducted to test which aspects of prosodic prominence facilitate word learning in native and nonnative perceptual language processing. As predicted, words under the narrow focus condition tended to be better recognized than words in broad focus condition.

The results obtained in the spliced conditions confirm findings of earlier studies (Akker & Cutler, 2003), as no effects could be established for the cues in the prosodic contour surrounding an accented word.

Results indicate partly different processing strategies for the two language settings L1 German – L2 English.

The probe detection task used in this experiment may have been a more accurate test of the influence of accent, target length and target position on committing a representation to memory than on accuracy and effectiveness of processing. We think that a phoneme detection task is well-suited to test the effectiveness of processing, and are, indeed, carrying out such an experiment.

Reference

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. Bilingualism: Language and Cognition, 6(2), 81-96.

Allan, D. (2001). Oxford placement test 2: Oxford University Press.

Barcroft, J., & VanPatten, B. (1997). Acoustic salience of grammatical forms: The effect of location, stress and boundedness on spanish l2 input

processing. In A. Perez-Leroux & W. R. Glass (Eds.), Contemporary perspectives on the acquisition of spanish: Production, processing, and comprehension (pp. 109-121). Somerville, Ma: Cascadilla Press.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. Perception & Psychophysics, 20, 55-60.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. Language and Speech, 40, 141-201.

Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. Cognition, 7, 49-59.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. Journal of Experimental Psychology: Human Perception & Performance, 14, 113-121.

Klein, W. (1992). Zweitspracherwerb. In (pp. 78ff.). Frankfurt/M.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for english sentences with prosodic cues. The Modern Language Journal, 84, 372-389.

Rooth, M. (1992). A theory of focus interpretation. Natural Language Semantics, 1.

Rooth, M. (1996). Focus. In S. Lappin (Ed.), The handbook of contemporary semantic theory. Oxford: Blackwell.

Anke Sennema, Ruben van de Vijver, Susanne E. Carroll, Anne Zimmer-Stahl
Universität Potsdam
SFB 632 „Informationsstruktur"
Postfach 601553
14415 Potsdam
Germany
sennema@rz.uni-potsdam.de

Structuring Information through Gesture and Intonation[*]

Stefanie Jannedy & Norma Mendoza-Denton

Humboldt Universität zu Berlin & University of Arizona

Face-to-face communication is multimodal. In unscripted spoken discourse we can observe the interaction of several "semiotic layers", modalities of information such as syntax, discourse structure, gesture, and intonation. We explore the role of gesture and intonation in structuring and aligning information in spoken discourse through a study of the co-occurrence of pitch accents and gestural apices. Metaphorical spatialization through gesture also plays a role in conveying the contextual relationships between the speaker, the government and other external forces in a naturally-occurring political speech setting.

Keywords: Gesture, Intonation, Spoken Discourse, Narrative Structure, Political Speech, Affect.

1 Introduction

It has been widely accepted that the gesture and intonation systems correlate, both aiding in the structuring of verbally rendered discourse (Cassell, 2000;

Loehr, 2004). Yet most of the studies on co-speech gesturing analyzed in the literature result from experimental elicitation and give rise to gestures that are narrative and/or descriptive in nature. That is, participants in laboratory settings are asked to describe something rather concrete such as fish-trapping constructions (Enfield, 2003), or to narrate preselected cartoon strips or films (McNeill, 1992). In this study, we analyze a video recorded sample from a corpus of spontaneous, naturally-occurring data gathered at public Congressional Town Hall Meetings (THMs) in Tucson, Arizona, in which a speaker engages in political discourse, a task much more abstract and goal-directed than elicited narrative, and one in which the speaker is trying to put a political viewpoint across. The speaker (whom we call by the pseudonym *Mary-Jane*) is a woman from Arizona who appears to be in her early forties; her primary interlocutor (U.S. House of Representatives Congressman Jim Kolbe, Republican, Arizona 5th district) is a middle-aged man in his late fifties. They stand in a constituent-representative relationship to each other in the United States political system. In this THM, the co-speech gesturing deployed by Mary-Jane is used not only to describe the political landscape as she sees it, but to persuade, cajole, shame, provide evidence and otherwise convince the interlocutor and the audience to adopt her point of view. Our findings indicate that gesturing not only correlates with grammar and supports the structuring of the information that is acoustically rendered; at the same time, use of the visuo-spatial field conveys information on the relationship posited by the speaker between the government, its constituents and outside forces.

Kendon (1996) thinks of gestures as a 'spill-over' effect from the effort of speaking. For him, gesture is a separate and distinct mode of expression with its own properties which can be brought into a cooperative relationship with spoken utterances and used in a complementary way. He suggests that a study on how phrases of gesture and phrases of speech are related would throw useful light on

how information is structured. Bolinger (1986, p. 199) proposes that "gesture and speech/intonation are a single form in two guises, one visible and the other audible", and argues that gesture and speech stem from the same semantic intent. McNeill et al. (2001, p. 23) argue that "the organization of discourse is inseparable from gesture and prosody" and that "[they are] different sides of a single mental communicative process" (cf. Enfield, 2003, p. 45-46). The results of McNeill's experimental studies indicate that motion, prosody and discourse structure are integrated at each moment of speaking.

Languages such as English or German, unlike for example Swahili (McNeill, 1992), use few morphosyntactic cues to structure discourse. Instead, information is structured primarily by syntactic and prosodic means. Vallduví and Engdahl (1996) discuss how information is packaged into prime units, and give evidence for the difference between English (which uses primarily intonation) and Catalan (using primarily syntax) in the packaging of information.

The role of gesturing during speaking and its role in the structuring of discourse has until recently been largely unexplored. Yet the relationship of gesture to speech increasingly captures the interest of human-computer interface designers aiming to model the movements of animated agents, since it has been shown that gestures facilitate information comprehension (Beattie & Shovelton 1999). Gesture researchers have further shown that subjects can recall information that was selectively presented in the gestural but not in the verbal channel (Kelly et al., 1999; Cassell et al., 1999; McNeill, 1992), all without being able to remember what channel the information was presented in. These results are provocative because they imply that as far as the design of animated agents is concerned, greater information density can be achieved by presenting part of the information acoustically and part of it (possibly complementary or reinforcing information) visually, bringing us closer to modelling the online

workings of face-to-face conversations. A thorough description of gestural-speech co-dynamics is thus necessary for an algorithmic approach to be formulated, and essential in the design of automated agents and in achieving some naturalness in their movements (Cassell, 2000 and references therein).

The purpose of this research is to describe and understand the timelines involved in a case study of the complementary presentation of information. Part of the information is presented through the speaker's verbal stream, while the broader and complementary political setting and the assumptions on which it rests are presented gesturally. We have independently verified that both of these meanings come across by asking audiences of graduate and undergraduate students in English-speaking universities (at the University of Arizona, the Ohio State University, and at University College Dublin in Ireland) to describe what they think the speaker was trying to get across. Three different modes of presentation of our data have been judged by the student populations:

1) Acoustic and visual information: Audiences who have seen the video and listened to the audio have repeatedly described the complementary information presented on both channels.

2) Acoustic information only: Audiences who have been presented with the audio portion describe the main assertions presented in the spoken discourse with no reference to the larger political landscape.

3) Visual information only: Audiences who have only seen the video without the audio merely impute anger and emotionality to the speaker.

These results highlight one well-known fact about the relationship of gesture when it is ancillary to spoken discourse: while spoken discourse has a high

referential resolution, that is, it is able to pick out referents with relatively little ambiguity, gesture has a low referential resolution, so most of the information presented gesturally is supportive to speech and not recoverable solely from the gestural channel (an obvious exception to this being gestural systems that are full-fledged languages, such as American Sign Language). Gestures and spoken discourse are thus complementary in nature, and have different affordances and limits to their ability to present information. Harper et al. (2000) have found that in a study of 3D multiplayer wargame interactions, gesture was used for much more than just simple deictic functions. They explain:

> …language facilitates complex queries with the ability to express quantification, attribute and object relations, negation, counterfactuals, categorization, ordering, and aggregate operations. Gesture is more natural for manipulating spatial properties of objects (size, shape, and placement) in graphical environments. (Harper et al. 2000:3)

In Harper et al.'s study, subjects participated in a wargame simulation which involved some players working around a scaled, 3D model of a battlefield, while other players were consulted remotely. In post-game discussions, participants remarked on the difficulty of communicating with those not working around the 3D model being used (2000:5). Because of their heavy reliance on definite descriptions in combination with gestures in the face-to-face game, players encountered difficulties when they were restricted to using only the verbal channel.

We hold there to be a parallelism between gesture and speech, both of them carrying meaning on at least three different planes: structure, content, and social meaning.

The acoustic signal is an enormously rich source of structured information. from a phonetic-phonological point of view, there are positional restrictions or co-occurrence restrictions of phonemes, the phonotactics that make up a specific language variety. There is a specific prosodic or rhythmic structure, such that some parts of a word are uttered louder and longer lending a specific syllable more perceptual prominence either to indicate lexical stress or informational salience. At the same time, obviously, the acoustic signal also contains other grammatical information. That is, lexical or semantic content is being transported by the choice or words spoken, and pragmatic content is chosen by the context the words are uttered in. Further, the acoustic signal transmits social content, that is, it also carries social information pertaining to the speaker (age, gender, ethnicity, etc.) and their interlocutors.

Co-speech gestures correlate with grammar, that is, they correlate with grammatical structure by indicating beats (correlated with pitch accents), or gestural phrases (correlated with syntactic phrases). These same gestures also carry semantic and pragmatic content by being deictic or emblematic or just emphatic beats on parts of the information emphasized by the speaker. Simultaneously, these same gestures also transmit social content in terms of how the speaker recounts and relates to the world or abstract universe surrounding them.

The gestural information stream is able to work online and incorporate context to take advantage of props and of the spatial location of the audience. We argue that the communicative constraints of the sociolinguistic situation in our case study maximized the need for the simultaneous presentation of information, since the speaker was only grudgingly given the floor in the first

place, was constrained in the time she was allotted for her turn at talk; while the interlocutor constantly threatened to interrupt her, attempted to shift his attention away from her and to cut her off. Thus, the sociolinguistic pressure was great for the speaker to pack as much information as possible into her short turn at talk. For our purposes, this results in a naturally-occurring situation where time and interactional constraints push the limits of both intonational and gestural information packaging.

2   Background

2.1   What do hand, arm, and mouth movements have in common?

Both the motor theory of speech production/perception (Liberman and Mattingly 1985) and the articulatory phonology account of speech as a stream of articulatory gestures (Browman and Goldstein, 1990, 1992) place articulatory movements produced during speech and "paralinguistic" gestures on the same cognitive plane, as they are both thought to be *coordinated patterns of goal-directed articulatory movements* (Tatham 1996). These approaches require that speech movements and gestural movements be accounted for by the same mechanisms, since it is assumed (contra Chomsky 2000, Fodor 1983, Marantz to appear) that the structures of language are not modular or unique in comprising a "language organ", but rather that these structures were derived and modelled on the pre-existing neural systems which had evolved for the control of body movement. Rizzolatti et al. (1988, 1999) posited a class of premotor neurons, the "mirror" neurons based on their experiments with monkeys. They explain: "…with this term [mirror neurons] we define neurons that discharge both when the monkey makes a particular action and when it observes another individual (monkey or human) making a similar action. […] Transcranial magnetic stimulation and positron emission tomography (PET) experiments suggest that a

mirror system for gesture recognition also exists in humans and includes Broca's area (1988:92)".

The results of these earlier studies of motor action have been replicated, and links shown between the sensorimotor system and the acoustic system both in monkeys and in humans (Berthoz, 1997: Fadiga & Craighero, 2003). The relationship between speech and limb movements is thus taken for granted, since both derive from elementary motor programs or *articulatory gestures*. According to Berthoz (1997:176) "Le mouvement est donc organisé à partir d'un répertoire de synergies qui compose autant d'actes possibles…une bibliothèque de mouvements facilement déclenchables. [*Movement is then organized beginning with a synergistic repertoire which makes up possible actions…making up a library of movements which are easily carried out.]*"

Both speech gestures (Browman & Goldstein, 1990) and hand/arm gestures then are finely coordinated and stem from higher level cognitive processes through which information is structured. While there is no one-to-one correspondence between form and meaning in gesturing, gestural organization maintains a tight link with the semantics of speech. That is, gestures strongly correlate with grammar and grammatical structures: it was found that the stroke (peak of effort) of a gesture occurs with the intonationally most prominent syllable of the aligned speech segment (Kendon, 1980; McNeill, 1992; Cassell, 2000; Loehr 2004).

The significance of the motor theory of action for our study is the following: we believe that when a speaker executes a particular gesture, the embodied nature of the gesture causes the interlocutor to generate an internal representation of the movements in order to decode and interpret the spatial field (concrete or abstract, see our discussion of Krifka 2005 below) depicted by the speaker. This is essentially a parallel to what the motor theory holds for speech (Liberman and Blumstein, 1990:147), where we practice "analysis by synthesis"

(Halle and Stevens 1959) in generating an internalized representation of the incoming signal.

Thus, given an interlocutor that shares spatial conventions, speakers may gesture and have their concrete and abstract meaning understood without uniqueness of referents. Spatial conventions exhibit cross-cultural variability: Haviland (1993), for example, has found that narrative retellings involving co-speech gesturing in the Gugu Yimidhirrh aboriginal language spoken in Queensland, Australia, exhibit absolute directionality, that is to say, when a speaker retells an incident their deictic pointing will refer to the same cardinal direction in which the relevant incident took place. Similar findings have been reported for Amerindian Languages such as Tzotzil (Brown and Levinson 1993) and Assiniboine (Farnell 1995). Other speakers, especially those from Western cultures, exhibit relative directionality, so that their gestures tend to be located relative to the ego (though some flexibility also exists here, with speakers able to present gestural situations from different points of view (McNeill 1992:190). There may additionally be abstract topic/comment structuring in the directions of pointing or in the handedness of a gesture (Krifka 2005). We hold that topic/comment structures will be present in preferential hand-gesturing, but will be forced to interact with culturally-conventionalized gesturing directionality. This would predict that, given the same story stimuli, speakers from absolute-directionality gesture traditions would encode topic/comment gestures differently from those of ego-centric gesture traditions. We exhort the next generation of linguists and linguistic anthropologists to conduct these experiments.

2.2   A typology of gestures

To gain a better understanding of the classification of gestures, a short digression into gesture theory seems necessary. From a kinetic point of view, gestural movements are described to have obligatorily three and at most five phases (McNeill, 1992): The *preparation phase* (optional) marks the beginning of the motion in which the parts of the body involved in the gesture leave the neutral position and move to the position necessary for the upcoming gesture. The *pre-stroke hold* (optional) is the position of the hand/arm at the end of the preparation phase before the beginning of the stroke. The *stroke* (obligatory) is the climax or peak of effort of the gesture. It is one of the most recognizable components of a gesture, it is synchronized with the linguistic forms, such as accented syllables it is co-expressive with. The *post-stroke hold* is the position the hand/arm remains in when the co-expressive spoken utterance is delayed. The *retraction* is the end of the motion in which the parts of the body involved in the gesture return to neutral positions. The *beat* gesture is smaller and is described to only have 2 phases which are typically flicks with the wrists or fingers. In the case of Mary Jane, we will see that her impassioned argument motivates her marking beats with her entire upper torso (Panel 4.2).

From a semantic or pragmatic point of view, gestures can have different attribute values:

1) *Deictic* gestures point towards a concrete or abstract referent;
2) *Beats*, sometimes called *batonic gestures* because they resemble a conductor keeping an orchestra in time, are rhythmic gestures that have no specific form but which are synchronized with the speaker's utterances.
3) *Iconic* gestures depict a concrete object or event in a narrative;
4) *Emblematic* gestures replace speech (for instance, shaking one's head to mean "no").

5) *Metaphoric* gestures represent an abstract idea.

McNeill (1992) calls the recurring combination of the same gestures with prosody and discourse organization *catchments*. These catchments are recognized from recurrences of gesture form features over a stretch of discourse (two or more gestures with partially or fully recurring features of shape, movement, space, orientation, dynamics etc.) and serve to offer clues to a cohesive linkage in the text in which it occurs.

Rather than assuming that gesturing only serves the interlocutor in structuring acoustically rendered information, we take it that gesturing aids both the speaker and the interlocutor: recall Harper et al.'s (2000) wargame simulations referenced above. It has also been observed that gestures occur no less frequently when talking over the telephone and the speaker cannot be seen by the listener (Cosnier 1982; Rimé 1982).

We also assume that the very same gestures that are so tightly aligned with grammar and prosody serve as a device to mediate between the speaker and the world. What does this mean? By having gestures that are ego-centered, and placing herself in the middle of a depiction of the political landscape, Mary-Jane sketches *through gestures* her view of an idealized public sphere, and of the relationship between a constituent and the broader powers of government.

2.3   Intonational Grammar

The grammar of intonation we are assuming goes back to Pierrehumbert (1980) and was formalized as a transcription system for prosodic annotation summarized in Beckman & Ayers (1994). According to this grammar of intonation there are pitch accents, intermediate phrases and intonation phrases. The grammar of intonation can be summarized as follows: each intonation

phrase (marked with '%' at the right edge) must contain at least one intermediate phrase (marked with a '-' at the right edge), which in turn must contain at least one pitch accent (marked with a '*').

Pitch accents are associated with the stressed syllables of words. These accents mark local prominences above the level of the word in an utterance. There is a fixed inventory of pitch accents, they can have different tonal shapes that are marked with labels such as H*, L* or L+H* etc. The '*' indicates that there is a pitch accent which is defined as a tonal target. The tonal events between pitch targets are accounted for by interpolation between these pitch targets. For example, one should observe a falling fundamental frequency contour when a H* tone target is followed by a L* tone target. Downstep is a phonologically triggered process by which a H* accent is downstepped (!H*) when it occurs subsequent to the downstep trigger. Often this trigger is a bitonal pitch accent type such as L+H* or L*+H accent. By definition though, the domain of downstep is the intermeditate phrase, thus, the trigger must occur in the same intermediate phrase as the downstepped accent. Downstepping is a compression of the pitch range lowering the F0 targets for the H* accents following the downstep trigger.

An intermediate phrase is a minor phrase and consists of one or more pitch accents plus a phrase accent associated with the right edge. This phrase accent can be either L or H. An intonational phrase consists of at least one or more intermediate phrases. The right edge of the intonational phrase has a phrase tone, taking the shape of L or H. These edge tones determine the shape of the F0 contour between the last pitch accent and the end of the phrase.

3   Data and Methods

We utilize video data collected in 2000-2001 for an ethnographic study of Congressional Town Hall Meetings (THMs) in Tucson, Arizona. The data forms part of a fieldwork project on political discourse, language and power. A total of approximately 20 hours of data was collected at locations around Southern Arizona and Washington, D.C. THMs are public fora announced in the media, on websites and through fliers to homes in the neighborhoods that are represented by Kolbe. A THM in this district normally lasts one and a half hours to two hours and is led by Congressman Kolbe, typically taking the form of an initial period of question-writing by the audience on slips of paper circulated by Kolbe staffers. This is followed by a rehearsed monologue from Rep. Kolbe in which he states as his aim the updating of his constituents on important happenings in Washington, D.C. Then Kolbe selects some of the questions that constituents have written out as questions to be answered without calling on anybody specifically. After this he might take a couple of spontaneous questions, run to the end of the allotted time and then invite those who wanted to talk to him further to stay and discuss matters after the official THM ends.

The video data of Rep. Kolbe and Tucson citizens that we analyze for this paper was recorded simultaneously with 2 video cameras, both located to the left of the audience (from their perspective), one pointed toward Rep. Kolbe and the other more aimed more generally toward the audience. Rep. Kolbe was miked with a lavaliere microphone plugged into camera 1, and the audience sounds was captured by a microphone mounted on camera 2. By aligning the sound tracks of the two videos, we were able to synchronize the videos exactly so as to gain accurate descriptions of the hand and arm movements from two different angles. Both camera 1 and camera 2 capture the audience from different vantage points.

On this particular occasion, on a Saturday morning in February 2001, the THM was held in the cafeteria of a midtown school. The congressman introduced the researchers and advised participants in the THM that they were being video taped for a research project and that their participation was strictly voluntary. If any of the constituents objected to being taped, they were encouraged to approach the researchers after the THM. None did.

The data we selected for microanalysis for this study lasts exactly 130 seconds (a reasonable amount of data within the gestural analysis literature (See Loehr 2004 for a discussion)), and was transcribed according to the ToBI intonational transcription framework in addition to being subjected to a modified McNeill-style gestural transcription. Eleven tiers of body movement and gestural transcription were coded.

3.1   ToBI and Gesture Transcriptions

Trained linguistics and anthropology students at the University of Arizona worked on transcribing the data on several different levels by using the program Praat. Orthographic transcriptions were done and then checked by another student. Prosodic transcriptions were made in a team of three students of linguistics by majority vote: They were trained to transcribe intonational events within the ToBI (tone and break index) framework based on the Pierrehumbert (1980) system of English intonation. In cases of uncertainty or disagreement, they discussed the issue until a consensus was reached. Then, all prosodic transcriptions were checked by an experienced labeler of US-English (one of the authors). It must be noted though that the data was very difficult to annotate due to the fact that the recordings were made in a naturalistic setting. In addition to room noises (random background noise) which shows up in spectrograms as energy in all frequency regions, the audio data we analyzed contains claps and

coughs from members of the audience, partial interruptions by the congressman and other noises that are unidentifiable. In figure 1 below, taken from our Praat display, we show the sound pressure wave (waveform) in the upper display, and a spectrogram (individual frequency bands) which was overlain with the fundamental frequency trace (F0) necessary for the tonal analysis. This is how we have coded lines 23-29 of the full transcription included as an appendix to this paper.
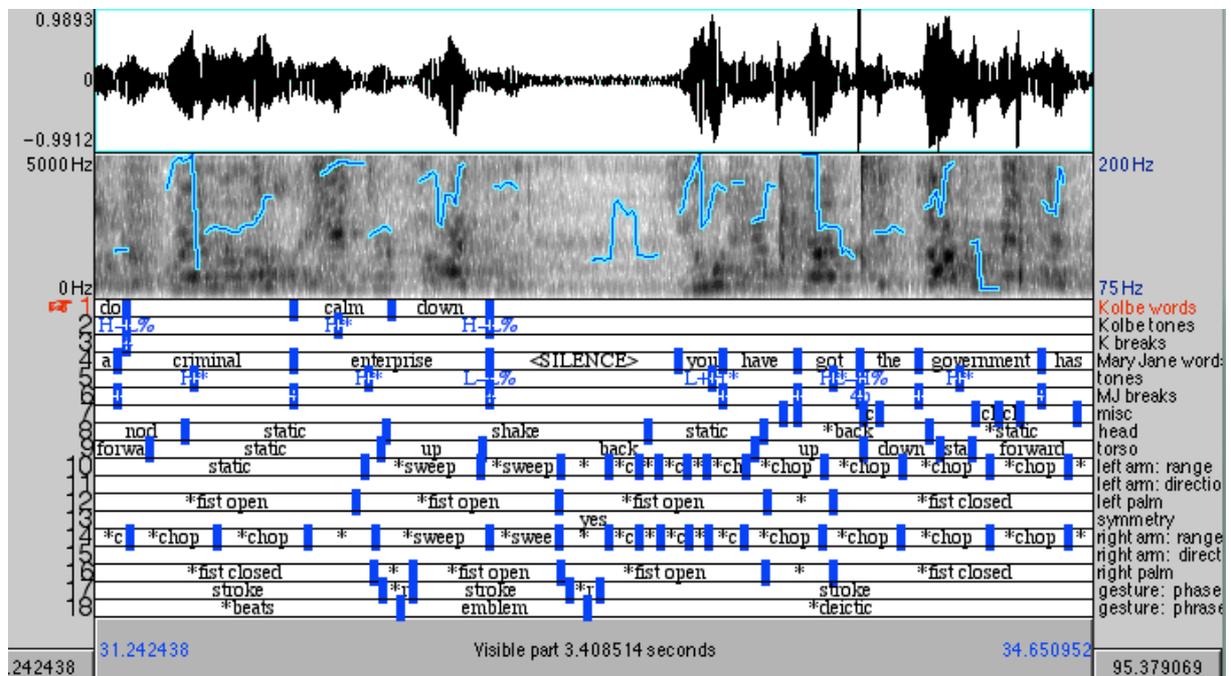

Figure 1: Praat display of multitier transcription of intonation and gesture.

Though there is a growing body of literature on gesture, no standard transcription system for manual gestures has been agreed upon and described. There are a couple of systems which are in relatively wide use; within language studies, the most prominent is McNeill's system, developed expressly for gestures. Based on gestural primes proposed by McNeill (1992), we transcribed movement of the left and right arm and the left and right hands on 6 tiers (range

of movement, direction of movement, palm configuration for each side). In addition, head and torso movements were transcribed as well as whether or not there was a symmetry in the movements of both hands and arms. We also annotated the gestural phases which have kinetic properties (preparation vs. stroke, for example) as well as the gestural phrases which are more semantic or pragmatic in nature (deictic vs. metaphoric etc.).

In order to have an exact alignment of audio and video, the data was looked at in ANVIL, a program that allows video and audio to be time aligned for transcription purposes. For the purpose of our study it became crucial not only to mark intervals during which gestural movements were performed, but to mark points as well. We were faced with the issue of data reduction since it seemed impossible to gain any insights from the amount of data available to us. As an initial step, we wanted to investigate if and when and how often pitch accents would co-occur with gestural movements. Since we had mainly coded intervals, and pitch accents are by default point events in time, we added another transcription level in which we coded what Loehr (2004) has called the *apex*.

The apex is the point in the hand or arm gesture during which (in Task Dynamics terms) the equilibrium of a particular gestural movement is reached (Browman & Goldstein, 1990). Task Dynamics holds that gestures are defined in terms of tasks (extend arm, point with finger etc.) which are the coordinated movements of several limbs. Further, the gestures are defined in terms of the dynamics that specify the motions in terms of a mass-spring model applied to articulation. Gestures are defined in terms of three specifications: 1. the stiffness of the gesture, the target of the movement and the phasing of the gestures with regard to each other. The stiffness roughly corresponds to the speed of the gesture, that is, speed is a reflection of the stiffness of the gesture. The stiffer the gesture, the faster it is being executed. When the target is being approached relatively slowly then there is less stiffness. The amplitude of the gesture is the

amount of displacement from the resting position. The phasing of gestures then is the intergestural timing, that is, the timing between gestures, specifying when one gesture begins and when the next gesture ends.

In terms of our specification of the apex, we adapt Loehr's (2004:89) definition. He describes the apex as the "peak of the peak" or as the "kinetic goal of the stroke". In task dynamics terms then this is the target of the gesture. The reason why this is described here in some detail is that we noticed that in more rapid, highly emotional speech, the co-speech gestures are not fully carried out, there often is just a succession of apices, tightly coordinated with pitch accents without the hands or arms returning to a resting position. The gestures overlap and cut each other off, only leaving the peaks of the gestures being observed. (In our example we find that the speaker nods her head for emphasis as well. We have however excluded this from our analysis and annotated hand, arm and body movements only).

4   Analysis

The total length of the sequence we selected for analysis is 2 minutes and 10 seconds. We have provided a full transcription of the whole sequence of Mary-Jane's monologue as Appendix A at the end of this paper. For microanalysis we have selected several segments which most clearly show the various gestural and intonational effects that we aim to illustrate. These segments are subdivisions created by us in our data and are not meant to be compared against each other, but to illustrate sequences that we felt had a unity of topic, sequencing and purpose within the total data set. They are presented in the order in which they occurred, as clusters of frame grabs in panels 1-4. Each panel has been divided into individual frame grabs in order to show the exact gestural sequences. Thus, P2:12-14 means Panel2, frames 12 through 14. We will follow

the following format in the presentation of this data: First we will present the transcript along with associated annotations, then the corresponding panels and finally our discussion and any illustrative figures related to them. Let's begin with the first sequence:

4.1   Panel 1 analysis: The drug war will be over.

M-J	the drug war will be over. **P1:1-3**
(0:20)	we won't have to ship guns and military to Colombia, **P1:4-11**
	it will be over. **P1:12-15**
	Bush won't have to talk to Fox about the drug war, **P1:16-24**
	it will be **P1:25-27**
	over. **P1:28-30**
Audience	[XXclappingXX]
	[there will not be a drug war] **P1:31-36**
	if you legalize drugs. **P1:37-42**

Panel 1: "It will be over"

In P1:5-11, Mary-Jane begins with a deictic gesture with the origin close to her chest. As she utters the phrase, "we won't have to ship guns and military to Colombia," she extends her right arm out and points out to the right, metaphorically placing Colombia to one side and away from the origin at ego.

This gesture has a clear parallel with later gestures, most notably the one in P1:16-24, where M-J does a bimanual pointing gesture immediately in front of her and then extends both arms to the left while saying, "Bush won't have to talk to Fox about the drug war", producing apices and pitch accents aligned with "Bush" and "Fox." Interestingly, though both Colombia and current Mexican president Vicente Fox (and by extension, Mexico) are located south of her, she has placed them to the right and left respectively, while Bush, who is north-east of her, has been placed directly in front of her (she herself is facing absolute south). Although we would not expect a speaker of English to exhibit absolute cardinal directionality, people in Tucson often do, since the city is two hours north of the border with Mexico and has dramatic mountain ranges which orient residents to absolute directions. This apparent disparity in the directions of well-known places was our first clue that something beyond simple deictic direction was being represented through Mary-Jane's gestures. A bird's eye representation of her gestures would show Bush, guns and military directly facing M-J, while Fox and Mexico would be to the left side and Colombia to the right (Figure 2). :
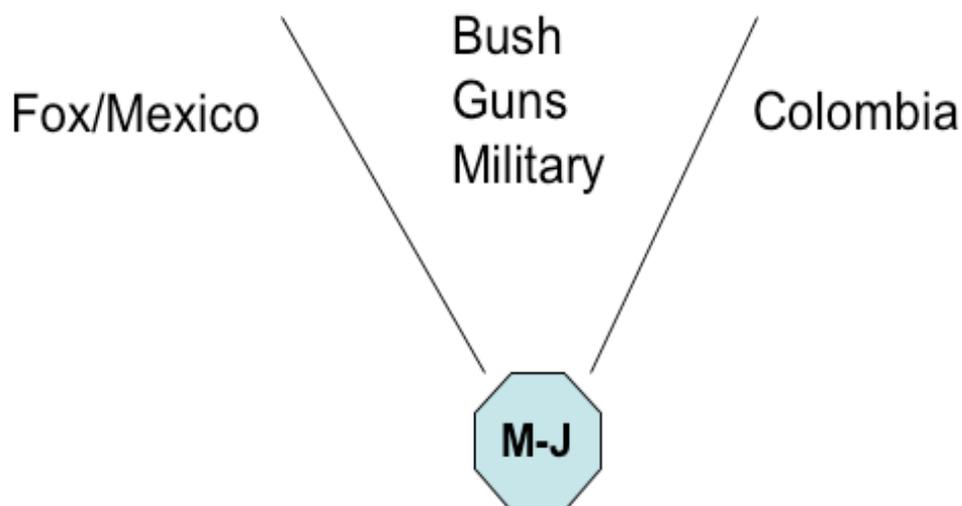

Figure 2: Birds' eye representation of spatialization of entities by M-J, Panel 1

An interesting catchment of gradually exaggerated gestures also occurs in this stretch. M-J repeats the phrase "it will be over" three times, each time with successively downstepped intonation and more emphatic gesturing. The first time, she brings her left hand (which is in ASL handshape G/X1 (McNeill 1992:88) in one downward stroke from eye level to shoulder level (this is visible in P1:1-4). For the second iteration of the phrase "it will be over," she draws a three-sided, open-bottom box with both hands, starting at the center top (P1:12-15; P1:14 shows both palms facing each other as she sweeps them down to make the sides of the box).

In the final iteration of the "it will be over" phrase (P1:25-30), she draws the complete box and slows down her speech, exhibiting an intonational phrase boundary and a gestural hold in the exact spot where she prepares to bring her hands in to close the box (P1:29), right before the last word, "over." The nested structural parallelisms appear thus:

1) lexical, intonational and syntactic parallelisms with coindexed pronouns at ACE;
2) intonational, gestural, and syntactic parallelisms at BD,
3) and repetition along with gestural parallelism and expansion at CE.

A   The drug war will be over

B   We won't have to send guns and military to Colombia

C   It will be over

D   Bush won't have to talk to Fox about the drug war

E   It will be over.

Tannen (1989) has argued that repetition in discourse has a cohesion and focussing function, serving to highlight information and structure listener

expectations. We believe the communicative constraints that we mentioned above motivate the use of four different semiotic levels of repetition (lexical, syntactic, phonological (intonation), and paralinguistic (gesture)), all within a space of nine seconds of spontaneous speech. The fine-grained coordination of moment-to-moment speech has been discussed in the conversation analytic literature, and to some extent in the literature that is called "emergent grammar," which is concerned with finding the emergent structure of speech as it happens in naturalistic interaction. Our description of the incorporation into emergent structure of affect-laden gesture and intonational phenomena contributes to this literature.

4.2   Panel 2 Analysis: Excuse me, hello?

	L+H* H* L-H%
M-J	because they're going to ex**pe**riment with **drugs**, (P2:1-3)
	L* L* L-H%
	like **ev**ery one has **done**, (P2:4-9)
	L* L-H% H* L-L%
	since **day** **one**. (P2:10-14)
Kolbe	ok,
	well I--[
	H* L-L%
M-J	] we ex**pe**riment. (P2:15-16)
	H* H* H* L-L%
	alcohol and to**ba**cco **kill** people (P2:17-23)

```
         H*              H*  H*              L-L%

    mariju**a**na doesn't **kill**   **a**nybody. (P2:24-28)

        L*+H              L-H%

    ex**cuse** me (P2:29-32)

    L+H*           !H-L%

    he**llo**: (P2:33-36)
```


Panel 2: "Excuse me, hello?"

This segment is annotated with the ToBI tones marking pitch accents (*) as well as phrase tones (-) and boundary tones (%). The bolded words indicate that there is a simultaneous occurrence of a pitch accent with a gestural apex. It is particularly noteworthy that Mary-Jane deliberately separates the phrase 'day one' into two intonational phrases. Both words are lent prominence by accenting them and by making pointing gestures where her right index finger lands on the upwards-open palm of her left hand (day: P2:10-11 and one P2:12-14). Note also that Mary-Jane produces the last gestural apex in this segment with her entire upper body as the articulator, which she abruptly stops in mid motion at a precarious 45-degree angle as she says to the Congressman "excuse me, hello".

The intonation contour L*+H L-H% occurring on 'excuse me' has been described as carrying a holistic pragmatic meaning ranging from uncertainty to incredulity (Ward and Hirschberg 1985, Hirschberg & Ward 1992). Hirschberg and Ward argue that "when speakers use the contour to express incredulity, they generally express that incredulity about a value already evoked in the discourse" (p. 243). A striking fact about this use of the uncertainty/incredulity contour is that it is not aligned with the utterance that one might expect it to, given the partially-ordered set relationship (poset) framework described in Ward and Hirschberg (1985). Consider the relevant part of the utterance once more:

       H*   H*   H*   L-L%

alcohol and to**bacco** **kill** people (P2:17-23)

         H*   H* H*   L-L%

mariju**a**na doesn't **kill**   **a**nybody. (P2:24-28)

       L+H*   !H-L%

ex**cuse** me (P2:29-32)

In denying that marijuana kills people, Mary-Jane invokes the possible set of substances that do kill people, and discursively selects for the interlocutor (Kolbe) and overhearers (other constituents) deadly substances that are *legal*. The incredulity being expressed here, signalled by the wide pitch range that the contour carries and which upgrades it from uncertainty (Hirschberg and Ward 1992), is in the assertion that it is the noxious substances that are legal while the innocuous substance is illegal. These discourse facts all line up with the licensed uses of this contour according to Ward and Hirschberg (1985), with one exception: the contour doesn't fall on the expected phrase (marijuana…), but instead on the *following* one (excuse me...). We believe that this finding has two interpretations with important implications:

1) It is possible that this is the point where a speaker under strong communicative demands finally reaches the limits of the processing capacity for online information alignment. In this short segment, Mary-Jane simultaneously aligns pitch accents with syllables contained in informationally prominent words, gestural apices with those pitch accents, while making a complex argument and taking parts of her body to the limits of their physical space. Could the canonical alignment of the incredulity contour have been given up in order to meet the extra processing demands of this task?

2) A second interpretation is that the speaker, knowing that she was going to use a conventionalized expression with its own L*+H L-H% contour ("excuse me, hello?" often is uttered this way colloquially in the United States), chose to suppress the first contour so as to avoid two identical contours together, following a kind of prosodic OCP, or presumably

because the meaning is already accessible from a single utterance of the contour.

On "hello", Mary-Jane produces a pitch contour that we described with the tonal sequence L+H* !H-L%. This contour has also been described as a 'calling contour' (Beckman & Ayers, 1994). It is the contour often used when shouting a name during the process of looking for somebody (for example to call somebody for dinner). It is our impression that Mary-Jane uses this contour to call Congressman Kolbe metaphorically to do his job, to reproach him.

4.3   Panel 3 Analysis: Alcohol and Tobacco

M-J	]why aren't you talking about alcohol and cigarettes? (P3.1:1-7)
	what you've done is you've taken that industry, (P3.1:8-12)
	and you are putting it on third world. (P3.1:13-17)
Kolbe	what--
M-J	they are consuming the tobacco now. (P3.1:18-23)
	that's where all the tobacco money is coming from. (P3.1:24-30)
Kolbe	[we're spending]
M-J	[they're still making] a profit on cigarettes (P3.2:31-37)
	cause they're selling them to all the [people] (P3.2:38-40)
	that we're dumping our chemicals on (P3.2:41-48)

Panel 3.1: "Alcohol and Tobacco."

Panel 3.2: "Alcohol and Tobacco."

The segment represented by panels 3.1 and 3.2 returns to the earlier metaphorical spatialization described in Figure 2 despite the intervening segment in panel 2. Also before panel 3, Mary-Jane has been referring to "kids you've been talking about today" who will go to prison because of their alcohol and drug use. Interestingly, the gesture in "kids" starts out directly in front of Mary-Jane, metaphorically in a space shared between the government and the people, and after drug use the "kids" get moved off to "prison," which also occupies the peripheral position earlier used for external entities (other countries) outside the relationship between the government and the people. Another movement from "us" to "them" takes place in Panels 3.1 and 3.2: as Mary-Jane discusses the practice of chemical dumping by the United States, she makes a gesture that starts out with both hands right behind her left shoulder (as though she were holding a ball and beginning to toss it, P3.1:10), and winds up gesturally in what is by now in her gesture system a stable place for the "third world," to the far right in her spatial system P3.1:18.
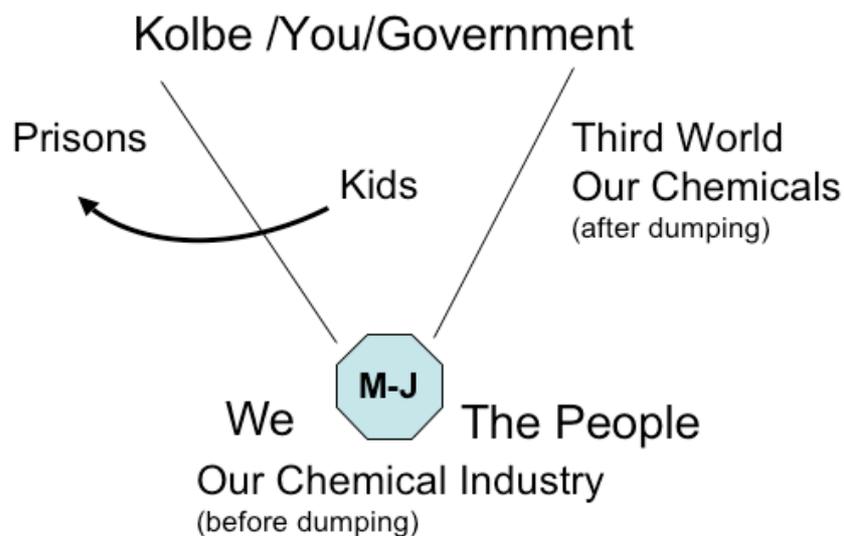

Figure 3: Bird's eye view of spatialized entities represented by M-J, Panels 3.1-3.2

Although a detailed discourse analytic account of the shifting pronominal referents in this data is beyond the aims and scope of this paper, we will note that the pronoun "we" variously refers to a) the people of the United States as a collective that produces chemicals that need to be dumped someplace ("people that we're dumping our chemicals on"); b) the United States as an entity that ships guns and military overseas ("we won't have to ship guns and military"); c) constituents that engage in collective behavior and do not share in the government's definition of what is legal and illegal ("we experiment"); and d) individual citizens who have rights to control their own bodies ("yes we do.") The pronoun "you" likewise shifts in reference to the addressee in general, to representative Kolbe, to current president G.W. Bush, and to the government in general. This variability in the instantiation of pronominal reference has been observed in political discourse cross-linguistically, as political figures (Van Dijk 2003, Wodak 1989) act on the political stage. In this case tracking the shifts in pronominal reference helps us to anchor and interpret the gestural data, and to make explicit the implicit political model that Mary-Jane sketches with her gestures.

4.4   Analysis Panel 4: My rights as a citizen

M-J   they want to make (P4.1:1)

(.)

hemp seed (P4.1:2-4)

a schedule one narcotic (P4.1:5-11)

(.)

when did I lose the right,(P1:12-17)

(.)

---

to eat, (P1:18-22)

(.)

the food, (P1:23-24)

(.)

I want, (P1:25-27)

(.)

to eat. (P1:28-30)

when did I lose my rights

as a citizen of this country, (P2:31-41)

to put into my body, (P2:42-45)

and listen to whatever, (P2:46-48)

or watch whatever I want. (P2:49-54)

In the sequence "to eat, the food, I want, to eat" is composed of 4 separate intonational phrases. Each of these phrases contains one intermediate phrase which in turn contains one pitch accent. There are accents on "eat", "food", "want" and "eat". We have observed that perceptually, these accents appear to occur in a downstepping relationship to each other (Liberman & Pierrehumbert, 1984). According to the tonal labeling conventions (ToBI) though, this would be infelicitous. Unfortunately, measurements of the fundamental frequency contour during the time point of the F0 maximum did not generate any insights because the signal was too perturbed. While the quality of the recording allows us to understand the message and hearing the pitch, the pitch tracking algorithms in different analyses software packages (Wavesurfer, Praat) logged values that were unreliable. We would like to point out though that we believe that there is a relationship between the pitch accents in this sequence and that the trigger of downstep can transcend the limits of an intermediate phrase.

Panel 4.1: "My rights as a citizen"

Panel 4.2: "My rights as a citizen."

The final panels for our gestural/intonational microanalysis are panels 4.1 and 4.2. In this segment Mary-Jane has taken advantage of a prop that she brought to the Town Hall meeting to dramatize her point: a hemp-seed chocolate chip cookie in a ziploc bag. She holds up the cookie and addresses the audience, showing them the offending item which is made with hemp seed (the seed of the marijuana plant). She attempts to highlight the absurdity of the government's decree that marijuana is illegal by claiming that she does not have the right to eat a chocolate chip cookie. She ends her performance (one of the issues for us is that we cannot be sure that this is not a rehearsed, prepared speech) with a dramatic flourish, by uttering two parallel constructions, with the same syntactic structure, in a slow and dramatically delivered style. They are presented below, in column format to highlight the parallelism. In conversation analytic transcription conventions, the periods in parenthesis mean that there were significant pauses in the speech stream.

1	2
when did I lose the right	when did I lose my rights
(.)	(.)
to eat	as a citizen of this country
(.)	(.)
the food	to put into my body
(.)	(.)
I want	and listen to whatever
(.)	(.)
to eat.	or watch whatever I want.

In this segment, not only is the syntactic/constructional parallelism evident above, but we further note that there occur to sequences of pitch accents in these

two utterances with downstep quality, despite the fact that in standard Tones and Breaks Indices theory it is thought to be the case that intonational downstep cannot happen across intonational phrase boundaries (Beckman & Ayers 1994). A downstepped H* accent is a type of high star accent, except that the tonal realization is being influenced by the accent preceding the high tone. Often a L+H* accent (a bitonal pitch accent) triggers downstep, so it is thought not to be able to occur across an intermediate phrase, because the trigger is displaced to the preceding intonational phrase. In this case, we would like to claim that for reasons of parallelism and emphasis, downstepped H* (notation: !H*) accents were used in places that have not been previously documented. These !H* accents coincide in our data with the apices of the gestures as well, since Mary-Jane rocks her entire body back and forth approximately 30 degrees to land at the apex of these movements on the !H* accents in both of these utterances.

This brings us to a more general discussion of the correlation between pitch accents and apices.

4.5   Pitch accents & Apices

In order to establish some kind of a measure of what the relation is between the occurrence of an accent and the occurrence of an apex, we counted the number of times Mary-Jane produced accents, that is, prosodic prominences, as well as the number of times where her gesturing included an apex, the peak of the gesture. The numbers are shown in the table below.

Segment	Dur in Sec	Apices	PA	Co-occurring Apices & PA		PA but no Apices	Apices but no PA
1. over	14.23	16 (100%)		16	(100%)		0 (100%)
			24 (100%)	16	(66.7%)	8 (33.3%)	
2. excuse me, hello	11.72	11 (100%)		11	(100%)		0 (100%)
			15 (100%)	11	(73.3%)	4 (26.7%)	
3. tobacco	14.26	16 (100%)		15	(93.8%)		1 (6.3%)
			27 (100%)	15	(55.6%)	12 (44.4%)	
4. lose the right	19.15	28 (100%)		26	(92.8%)		2 (7.1%)
			32 (100%)	26	(81.3%)	6 (18.8%)	
Total	59.36	71 (100%)		68	(95.7%)		3 (4.2%)
			98 (100%)	68	(69.4%)	30 (30.6%)	

Table 1: "(Co-)occurrence of Pitch Accents and Gestural Apices"

The general tendency apparent from this table comparing columns 'Apices' and 'PA' is that there are always more pitch accents than apices in any given segment. Note that we are not comparing the segments with each other, we have merely stated the numbers separately as to make the counting more transparent and easier. As we can see from column 'PA but no Apices', in verbally rendered speech, information is highlighted by acoustic prominences that obviously must not have visually co-occurring gestural apices. On the other hand, it only rarely happens that there is an apex occurring without a pitch accent (see column 'Apices but no PA'). This suggests to us either that some prosodic prominences are just rhythmic in nature or that not all prominences are semantically 'worthy' of being marked by co-speech gestures. It is possible also that the information given on these two different planes is complementary in nature, highlighting *different parts* of the message.

In column, 'Co-occurring Apices & PA', the numbers in the unshaded box (e.g. 16 (100%)) indicates that whenever there is a gestural apex, there also is a pitch accent. The number in the shaded box indicates that only 66.7% of all pitch accents were accompanied by an apex, a beat gesture. In terms of the total

numbers, we find that 95.7% of all apices were accompanied by a pitch accent whereas only 69.4% of all pitch accents were additionally marked by a gestural apex.

We have looked for the co-occurrence of accent location and gestural apex. However, it may be promising also to differentiate between the accent types, which we have not done so far. It is possible that certain (less prominent) accent types such as a downstepped !H* or a L* may be reinforced by gestural means. However we can make no claims regarding this hypothesis.

5 Discussion

We found that speech and gesturing are two different channels/modes of information transfer which allow for different content to be transmitted. If we assume the validity of Bolinger's (1986) claim that "gesture and speech stem from the same semantic intent […]" then we commit ourselves to the notion that some degree of preplanning is involved in generating not only speech output but also gestural output in order to convey information on different planes. How information is structured and divided up across the two channels is not understood at this point. From our data it appears that complimentary and contextual information is transmitted via gestures while concrete assertions are made explicit via speech. We also do not know what constraints exist on (pre-) planning complex gestures that we know are time aligned with linguistic structure in the final output.

Gestures are not just involuntary movements but finely coordinated structures of motion, aligned with semantic content. Since speech can be understood over the telephone or in the dark (with a complete absence of visual cues), we must assume that gestures facilitate the information transfer from the speaker to the listener/viewer but are not necessary for successful transmittal of

content. Gestures play an important role for the naturalness of speech and for cuing speaker stance. On the other hand, it is also known that speakers gesture while speaking when nobody is there to see them. Therefore, it appears that gesturing is not just a facilitation device for the listener/viewer but a mode of self expression for the speaker.

Gesturing exhibits cross-cultural differences, and it is, just as other communicative actions, learned behavior. The mechanisms of acquiring co-speech gesturing consists of coordinating the individual tasks of the complex gestures with each other (for example, lift right arm, rotate palm upward, release arm in this constellation) and then coordinating these motions with speech so that points of informational prominence in speech are accompanied for example by apices in gesture.

An infant as young as a few hours displays the ability to mimic the sticking out of one's tongue when prompted (Meltzoff & Moore 1977, p. 78). This suggests that our cognitive systems provides for learning by example and imitation. The task is a formidable one, as inverse mapping has to take place: the infant observes tongue movement via the visual channel, and then has to map the observed movements to own motor patterns of the articulators (opening lips, lower jaw, extending tongue) in order to perform the task of sticking out the tongue.

It appears that the same type of mechanisms should be involved to learn other motor skills such as finely coordinating hand, arm and body movements to be timed with speech: based on our casual observation, often the offset of a complex gesture co-occurs with the end of a prosodic phrase (intermediate or intonational phrase) or as we have shown in this paper, apices, the peaks of the gestures, co-occur with pitch accents. An interesting test case would be an investigation of the type and timing of co-speech gesturing of people who have been blind since birth.

If the interpretation of intonation contours, that is, if the interpretation of prosodic focus is partly determined by context, we have to take into account that this context is not just provided verbally or relates back to knowledge the interlocutors possess already. Rather, in face-to-face communication the gestural channel is able to provide the speaker's stance, or part of the context in which to interpret the utterance. It appears though that the verbal and gestural channels are interpreted simultaneously and holistically so that the semantic content can be recalled but the presentation and structure of information is more fleeting in nature and thus cannot easily be teased out by the receiver. That is, we tend to remember meaning rather than form but also make inferences which let listeners arrive at an interpretation (Bransford, Barclay, and Franks 1972).

6   Conclusion

In this work we have carried out a case study of the coordination of spontaneous speech with gesture, focusing on intonational alignment with pitch peaks (after Loehr 2004), and have found that wherever there is a gestural apex in our data, there is also a pitch accent. The reverse, however, is not true, because pitch accents often occur without gestures or the apices are phase-shifted from the gestural peaks. Our results concur with those of Loehr (2004) for laboratory data, and indicate that gestural phenomena are in robust co-occurrence with pitch accents in both laboratory and spontaneous speech. Our findings also include the discovery of a downstep relationship across intonational phrase boundaries, as well as the pervasive use of lexical, syntactic, gestural, and intonational parallelism in the performance of a speaker under high pressure in an affect-laden, spontaneous communicative situation. While the intonational and gestural alignments were observed within and across intonational phrases, metaphorical gestural spatialization had a larger domain, it took longer to

unfold, and sketched out complementary information on this speaker's notion of the relationship between a political representative, the government, and outside entities.

7   Appendix: Transcription of Town Hall Meeting

Town Hall meeting at St. Cyril's, midtown Tucson school, February 2001. Filmed with Sony DVTR8 mini-DV camera, sound aligned with audience microphone, quicktime clip name: kolbestcyril2(2/18,DVTR8).mov

Kolbe   and since you wanna comment on [this

               [I'm sorry]

               We're gonna] get your comment[

 M-J   ]my--

               my only problem with--

(0:04)   n- n- n- not legalizing all drugs?

Green   right.

M-J   you take the criminal element out of it,

               when you end prohibition,

               (.)

               this is what we saw in the thirties.

               the drug war will be over.

(0:20)   we won't have to ship guns and military to Colombia,

               it will be over.

               Bush won't have to talk to Fox about the drug war,

               it will be

               (.)

               over.

Aud	[xxclappingxx]
	[there will not BE a drug war]
	if you legalize drugs.
Kolbe	ok [[calm down,
(0:30)	calm down,
	(.)
	calm down,]]
M-J	[[you are making it a criminal]] enterprise,
Aud:	boooo]
	[you have got-]
Aud:	[boooo]
	the government has made it a criminal enterprise,
	the government is making money,
	you are making money hand over fist,
(0:40)	you are building prisons so fast it's disgusting,
	you're not putting any of that money into education,
	you're locking those-
	you're building prisons to lock those kids up that you're talking about today,
	they're going to go to prison.
(0:50)	because they're going to experiment with drugs,
	like every one has done
	Since day one.
Kolbe	ok,
	well I--[
M-J	]we experiment.
	alcohol and tobacco kill people.
	marijuana doesn't kill anybody.

excuse me,

hello-o:

(.)

Kolbe   ok.

      (.)

      well--[

M-J   ]why aren't you talking about alcohol and cigarettes?

what you've done is you've taken that industry,

and you are putting it on third world.

Kolbe   what--

M-J   they are consuming the tobacco now.

that's where all the tobacco money is coming from.

Kolbe   [we're spending]

M-J   [they're still making] a profit on cigarettes cause they're selling

them to all the [people]

that we're dumping our chemicals on

Kolbe   [we're--]

we're spending a lot of money on a--

(.)

on education,

which is what I think—

(.)

on tobacco,

which is what I think we need to be doing,

(.)

[spending a lot on that].

M-J   [proper name] wants to make this

a schedule one narcotic.

A chocolate chip cookie,

ok?

a chocolate chip cookie,

this is what they tried to run through on October the thirtieth,

they want to make

(.)

hemp seed

a schedule one narcotic.

when did I lose the right,

(.)

to eat,

(.)

the food,

(.)

I want,

(.)

to eat.

(1:40)   when did I lose my rights

(.)

as a citizen of this country

(.)

to put in my body

(.)

and listen to whatever

(.)

or watch whatever I want.

Kolbe   well,

(.)

	uh,
	there--
	there always have been limits on doing--
	on some things.
	you do—
	we do--
	you do not have[
M-J	]I limit myself,
	[you don't limit me thank you].
Kolbe	[y:--
	y:--
	you] don't have absolute rights to do everything.
M-J	yes we do.
Kolbe	and we do--
	(.)
	we never have had.

References

Beattie, G., & Shovelton, H. (1999) An Experimental Investigation of the Role of Iconic Gestures in Lexical Access Using the Tip of the Tongue Phenomenon. *British Journal of Social Psychology*, 90, pp. 35-56.

Beckman, M. E. & Ayers, G. E. (1994) *Guidelines to ToBI Labelling (version 2.0)*. The Ohio State University.

Berthoz, A. (1997) Le Sens du Mouvement. Paris: Editions Odile Jacob.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972) Sentence Memory: a Constructive Versus Interpretive Approach. *Cognitive Psychology* 3, pp. 193-209.

Bohman, J. (1996) *Public Deliberation.* Cambridge, MA: MIT Press.

Bollinger, D. D. (1986) Intonation and Gesture. *American Speech*, 58.2, pp. 156-174.

Brown, P., & Levinson, S. C. (1993) "Uphill" and "Downhill" in Tzeltal. *Journal of Linguistic Anthropology* 3(1), pp. 46-74.

Browman, C. P., & Goldstein, L. (1990) Tiers in articulatory phonology, with some implications for casual speech. In T. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, pp. 341-376. Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992) Articulatory Phonology: An overview. *Phonetica,* 49, pp. 155-180.

Cassell, J., McNeill, D., & McCollough, K. E. (1999) Speech-gesture mismatches: Evidence for One Underlying Representation of Linguistic and Nonlinguistic Information. *Pragmatics and Cognition*, 7(1), pp. 1-33.

Cassell, J. (2000) A framework for Gesture Generation and Interpretation. In R. Cipolla and A. Pentland (eds.) Computer Vision in Human-Machine Interaction. Cambridge: Cambridge University Press. http://www.soc.northwestern.edu/justine/publications/gesture_wkshop.pdf

Chomsky, N. (2000) Minimalist Inquiries: The Framework. In *Step by Step*, R. Martin, D. Michaels, and J. Uriagereka (eds.), pp. 89-155. Cambridge, MA: MIT Press.

Cosnier, J. (1982) Communications et langages gestuels. In J. Cosnier, A. Berrendonner, J. Coulon, and C. Orecchioni (eds.) *Les voies du langage: Communications verbales, gestuelles, et animales*, pp. 255-304. Paris: Dunod.

Enfield, N. (2003) Producing and Editing Diagrams Using Co-Speech Gesture: Spatializing Non-spatial Relations in Explanations of Kinship in Laos. *Journal of Linguistic Anthropology* 13, pp. 7-50.

Ekman, P. & W. Friesen (1972) Hand Movements. *The Journal of Communication* 22, pp. 353-374.

Farnell, B. (1995) Human Action Signs in Cultural Context: the Visible and the Invisible in Movement and Dance. Metuchen, NJ: Scarecrow Press

Fadiga, L. & Craighero, L. (2003) New Insights on Sensorimotor Integration: From hand action to speech perception. *Brain and Cognition* 53, pp. 514-534.

Fodor, J. (1983) *The Modularity of Mind.* Cambridge, MA: MIT Press.

Halle, M. & Stevens, K. (1959) Analysis by Synthesis. In W. Wathen-Dunn and L.E. Woods (eds.) *Proceedings of the Seminar on Speech Compression and Processing.* AFCRC-TR-59-198, December 1959, Vol. II, paper D7.

Harper, L., Loehr, D. & Bigbee, A. (2000) Gesture is not just Pointing. Presented at and published by the International Conference on Natural Language Generation.

Haviland, J. B. (1993) Anchoring, Iconicity, and Orientation in Guugu Yimidhirr Pointing Gestures. *Journal of Linguistic Anthropology*, Vol. III(1), pp. 3-45.

Hirschberg, J. & Ward, G. (1992) The Influence of Pitch Range, Duration, Amplitude and Spectral Features on the Interpretation of the rise-fall-rise Intonation Contour in English. *Journal of Phonetics*, 20, pp. 241-251.

Goodwin, C. (1981) *Conversational Organization. Interaction between Speakers and Hearers.* New York: Academic Press.

Kelly, S., Barr, D., Church, R., & Lynch, K. (1999) Offering a Hand to Pragmatic Understanding: The Role of Speech and Gesture in Comprehension and Memory. *Journal of Memory and Language,* 40, pp. 577-592.

Kendon, A. (1972) Some Relationships between Body Motion and Speech. In Siegman, A. & Pope, B., (eds.) *Studies in Dyadic Communication*, pp. 177-210. New York: Pergamon Press.

Kendon, A. (1980) Features of the Structural Analysis of Human Communicational Behaviour. In von Raffler-Engel, W. (ed.), *Apsects of Nonverbal Communication*. Swets & Zeitlinger, Lisse, Netherlands, pp. 29.43.

Kendon, A. (1996) Gesture in Language Acquisition. *Multilingua* 15, pp. 201-214

Krifka, M. (2005) A Functional Similarity between Bimanual Coordination and Topic/Comment Structure. Paper presented at the Blankensee Colloquium,

Berlin Schmöckwitz: Language Evolution: Cognitive and Cultural Factors, July 14-16, 2005.

Liberman, P. & Blumstein, S. (1990) Speech Physiology, Speech Perception, and Acoustic Phonetics. *Cambridge Studies in Speech Science and Communication*. Cambridge: Cambridge University Press.

Liberman, A. & Mattingly, I. (1985) The Motor Theory Revised. *Cognition* 21, pp. 1-36.

Liberman, M. & Pierrehumbert, J. (1984) Intonational Invariants under Changes in Pitch Range and Length. In M. Aronoff & R. Oehrle (eds.) Language Sound Structure. MIT Press, Cambridge, MA.

Loehr, D. (2004) *Gesture and Intonation.* Doctoral Dissertation, Georgetown University, Washington, DC.

Marantz, A. (to appear) Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review.* Accessible from author's website: http://web.mit.edu/linguistics/www/marantz/index.html

McNeill, D. (1992) *Hand and mind.* The University of Chicago Press.

McNeill, D. (2000) *Language and gesture.* Cambridge: Cambridge University Press.

McNeill, D., Quek, F., McCullough, K.-E., Duncan, S., Furuyama, N., Bryll, R., Ma, X.-F. & Ansari, R. (2001) Catchments, Prosody and Discourse. *Gesture* 1:1, pp. 9-33.

Meltzoff, A. N., & Moore, M. K. (1977) Imitation of Facial and Manual Gestures by Human Neonates. Science, 198, pp. 75-78.

Pierrehumbert, J. (1980) *The Phonology and Phonetics of English Intonation.* Doctoral Dissertation, MIT.

Pierrehumbert, J. & Hirschberg, J. (1990) "The Meaning of Intonation Contours in the Interpretation of Discourse." In Cohen, Morgan, and Pollack, (eds.) *Plans and Intentions in Communication and Discourse.* MIT Press, pp. 271-311.

Rimé, B. (1982) The Elimination of Visible Behavior from Social Interactions: Effects on Verbal, Nonverbal, and Interpersonal Variables. *European Journal of Social Psychology* 12, pp. 113-29.

Rizzolatti G & M. Gentilucci (1988) Motor and Visual-Motor Functions of the Premotor Cortex. In Rakic, P. and W. Singer, eds. *Neurobiology of the Neo Cortex*. Dahlem Workshop Reports. Chichester: John Wiley.

Rizzolatti, G., Fadiga, L., Fogassi, L., Gallese, V. (1999) Resonance Behaviors and Mirror Neurons. *Arch Ital Biol.* May 137(2-3), pp. 85-100.

Tannen, D. (1989) Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse. Cambridge: Cambridge University Press.

Tatham, M. (1996) Articulatory Phonology, Task Dynamics and Computational Adequacy. *Proceedings. Institute of Acoustics.* 18.

Vallduví, E. & E. Engdahl (1996) The Linguistic Realization of Information Packaging. *Linguistics,* 34, 3(343), pp. 459-519.

Van Dijk, T. A. (2003) Text and Context of Parliamentary Debates. In Paul Bayley (ed.), Cross-Cultural Perspectives on Parliamentary Discourse, pp. 339-372. Amsterdam: Benjamins.

Ward, G. & J. Hirschberg (1985) Implicating Uncertainty: the Pragmatics of Fall-Rise Intonation. *Language,* 61, pp. 747-776.

Wodak, R. (ed.) (1989) Language, Power, and Ideology: Studies in Political Discourse. Amsterdam Philadelphia: J. Benjamins Co.

Stefanie Jannedy
Humboldt Universität zu Berlin
SFB 632 „Informationsstruktur"
Mohrenstr.
14415 Berlin
Germany
jannedy@ling.ohio-state.edu