ISIS Vol. 1

2004
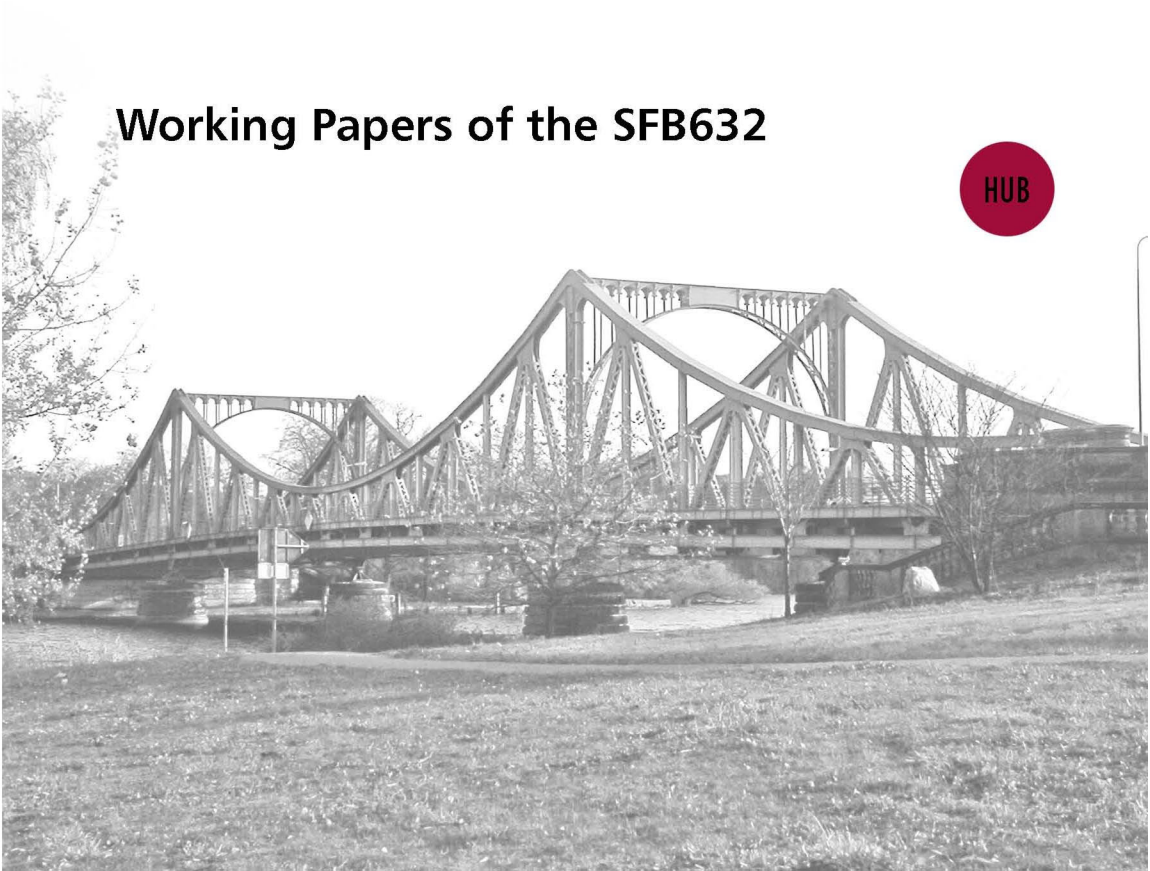
Interdisciplinary Studies on Information Structure Vol. 1

Working Papers of the SFB632

UP

HUB

Interdisciplinary
Studies on Information Structure

Preface

We are pleased to present the first volume of *Interdisciplinary Studies on Information Structure (ISIS)*. ISIS is the working papers series of the SFB 632 "Information structure." SFB 632 is a research group, funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), that brings together scientists from the areas of linguistics, psychology and English studies at the University of Potsdam, and linguistics and African studies at the Humboldt University Berlin. SFB 632 consists of 13 projects, all of which aim to formulate integrative models of information structure from different perspectives.[1]

This first volume includes 7 working papers from 5 different projects. Three papers from the A1 project by Gisbert Fanselow, Caroline Féry and Laura Herbst, and Shinichiro Ishihara focus on syntax-phonology interface issues. A paper by Cornelia Endriss and Stefan Hinterwimmer from the A2 project discusses the semantic problem of adverbial quantification. A paper by Ekaterina Jasinskaya, Jörg Mayer, and David Schlangen from the A3 project surveys the interface between discourse structure and information structure. A paper by Katharina Hartmann and Malte Zimmermann from the B2 project examines focus strategies in Tangale, a Chadic language spoken in Africa. A paper by Stefanie Dipper, Michael Götze, Manfred Stede, and Tillmann Wegst from the D1 project introduces ANNIS, the database for all the projects in SFB 632, which the D1 project is currently developing.

Although this volume does not include papers from all the projects, more papers will appear in future volumes. We will publish, through the ISIS series, working papers from our projects, as well as proceedings of the conferences/workshops organized by SFB 632.
With papers from such variety of research fields discussing a common topic, we hope that this volume (and moreover, the whole ISIS series) will interest and stimulate many scientists working on information structure.

<div align="right">

Shinichiro Ishiahra
Michaela Schmitz
Anne Schwarz

</div>

---

[1]Please visit the website for more information about SFB 632:
`http://www.ling.uni-potsdam.de/sfb/`

Contents

Cyclic Phonology–Syntax-Interaction:
Movement to First Position in German[*]

Gisbert Fanselow

Universität Potsdam

This paper investigates the nature of the attraction of XPs to clause-initial position in German (and other languages). It argues that there are two different types of preposing. First, an XP can move when it is attracted by an EPP-like feature of Comp. Comp can, however, also attract elements that bear the formal marker of some semantic or pragmatic (information theoretic) function. This second type of movement is driven by the attraction of a *formal* property of the moved element. It has often been misanalysed as "operator" movement in the past.

Keywords: topicalization, focus movement, operator movement, A-bar-movement

1   Introduction and Overview

For more than two decades, (most) syntacticians took it for granted that syntax and phonology interact in a global way: phonological rules apply as a block to the output of a *complete* overt syntax derivation (Chomsky 1981). However, in the context of the overall shift from representational to derivational models

---

(Chomsky 1993, 1995), it seemed natural to assume that the spellout operation (interpreting abstract syntactic structures morphologically and phonologically) is cyclic itself (Chomsky 2000). In this respect, (minimalist) syntax takes up suggestions made more than thirty years ago by Bierwisch (1968) and by Bresnan (1972). Bresnan argued that the cyclic nature of the syntax-phonology interaction was evident even when one considered simple wh-questions.

(1a) exemplifies the effects of the *Nuclear Stress Rule* (NSR), that places primary stress on the rightmost/most deeply embedded element in VP. In (1b), primary stress falls on an element that is not part of VP. Still, stress placement in (1b) is not in conflict with the NSR, Bresnan argues: the NSR places the primary accent on *what books* when that phrase is still in VP. If the application of the NSR precedes wh-movement, (1b) has a straightforward explanation.

(1)   a.   John said that Helen had written this BOOK

      b.   John asked what BOOKS Helen had written

Phonological rules and syntactic rules thus interact cyclically. If phonological rules may be applied prior to Move α, details of the movement operation should depend on the outcome of phonological rules. This paper argues that this expectation is borne out, e.g., by data as simple as German (2), when it answers questions such as *what happened last weekend?* Elements that can be fronted in a VP- or IP-focus utterance are identified by their phonological properties. The phenomenon that phonological properties determine which categories move is even more widespread, both in terms of constructions, and languages.

(2)   Ein BUCH hab   ich   gelesen
       A   book   have I.NOM   read
       "I read a book"

Our argument is framed in a minimalist syntax[1], in which the need of checking (formal) features of functional heads triggers the movement of XPs and Xs. We will sketch a model of movement to first position in German main clauses in section 2: XPs move to first position either because they are attracted by an operator feature (an assumption that we revise later), or they do so in the context of a "stylistic fronting" operation placing the leftmost element in IP in front of the finite verb.

In case of operator movement, attraction can be confined to part of the operator phrase. The basic data supporting such a *pars pro toto movement* (ppt-movement) will be introduced in section 3, where we will also show that the category undergoing ppt-movement is picked on the basis of its phonological properties. Operator movement and ppt-movement can be unified if the feature that is attracted in German verb second constructions encodes a formal (phonological, morphological) rather than semantic-pragmatic ("topichood") property. Section 4 presents some data that show that ppt-movement is not confined to German, while section 5 introduces ppt-movement data with a more complex information structure. In section 6, we compare the ppt-movement approach with remnant movement analyses.

2   Two Types of German Main Clauses

German main clauses involve at least two movements: the preposing of the finite V, and the subsequent placement of some XP in front of it, see Thiersch (1978), den Besten (1989), Vikner (1995), among many others.

(3) a. [Ich [gestern   [[ein Buch]   gelesen]] hab] →
          I   yesterday   a   book   read   have

---

[1]   Our basic argument is also valid in all models in which movement operations must be licensed, i.e., also in OT syntax models.

    b.   [Hab [ich [gestern [[ein Buch] gelesen]] ~~hab~~]] →

    c.   [Ein Buch [hab [ich [gestern [[~~ein Buch~~] gelesen]]]]]]
        "I read a book yesterday"

Den Besten (1989) argues that the finite V moves to C in (3b). Identifying the best theory of verb movement turns out to be surprisingly difficult (see Zwart 2001, Fanselow 2003, G. Müller 2003, Nilsen 2003), but details need not concern us since our conclusions are independent of them. Let us therefore simply assume that the finite V goes to C in verb second main clauses. The placement of some XP to the left of V in C can then be understood as movement to the specifier of CP.

    The early literature on German verb second clauses largely ignored the fact that there is little arbitrariness in the identification of the element that is placed into Spec,CP. There are rules to be followed, a fact that is, however, slightly obscured by the existence of two classes of main clauses in German.

2.1 Main clause type A: attraction of an operator

Constituent questions such as (4) constitute the most straighforward example of the first class of German main clauses: an operator moves to Spec,CP.

(4)   [$_{CP}$ Was [$_{Comp}$ hat] [$_{IP}$ er gelesen?]]
           what   has   he read

The following analysis (in line with the wh-criterion, Rizzi 1991) suggests itself: C possesses an EPP-like feature that triggers the attraction of some constituent (see Chomsky 1998), but only those categories are attracted that match further featural specifications of C. Thus, when C has a [+wh]-feature because it heads a constituent question, only wh-phrases will be attracted.

[+wh] is not the only attracting operator feature. (5) is an answer to (4). Focus XPs are attracted by C when C bears a [+foc]-specification. Unlike wh-phrases, focus XPs do not have to move. (6) is as good an answer to (4) as (5).

(5) [$_{CP}$ Ein BUCH [hat [$_{IP}$ er gelesen]]]
       a   book   has   he read
       "He has read a book"

(6)   Er hat ein BUCH gelesen

(7) suggests that C may also have a [+top]-feature: topics can show up clause-initially. According to Frey (2004), topic phrase have to be preposed (but see Fanselow 2003), but they may also land in positions below C.

(7)   (Soll ich was über Hans erzählen? "Do you want me to say something
          about Hans?")

          Diesen   Verbrecher hat man   endlich verhaften können
          this.ACC   criminal   has one   finally arrest   could
          "One has finally been able to arrest this criminal"

Operator attraction often involves the pied piping of larger categories, as in (8).

(8) a.   An wessen Schwester hat er   einen Brief geschrieben
          at whose   sister   has he.NOM a.ACC letter written
          "whose sister has he written a letter to?"

     b.   An SABINES Schwester hat er einen Brief geschrieben
          at Sabine's sister   has he a   letter written
          "He wrote a letter to SABINE's sister"

2.2 Main clause type B: Stylistic Fronting

In many main clauses of German, the initial element is neither a wh-operator nor

a topic or a focus. The need for distinguishing a second class of German main clauses was first recognized by Travis (1984). Although many of her arguments may have turned out to be incorrect (see, e.g., Gärtner & Steinbach 2003), it cannot be denied that subject-initial sentences follow conditions of use other than those of the sentences discussed above. E.g., the former can be thetic, uttered in out of the blue contexts. Similarly, subjects may appear in first position when a focus phrase is not in Spec,CP, as (6) illustrates. Subjects may *always* appear clause-initially without being a topic or a focus.

(9)   Was ist geschehen? "What happened?"

Ein Kind   hat seinen   Schlüssel verloren
A   child has his   key   lost
"A child lost his key"

The analysis of this construction is somewhat obscured by the fact that elements other than subjects can show up in clause initial position without special pragmatic force. Thus, the examples in (10) can be thetic: dative arguments of unaccusative and passive constructions (see Lenerz 1977), and sentential (Koster 1978) and temporal (Frey 2000) adverbs are like subjects in that they can be clause-initial without being a topic or focus (but see also Jacobs 2001).

(10) a.   Einem   Studenten ist ein   Fehler   aufgefallen
          a.DAT   student   is   a.NOM mistake   struck
          "A student noticed a mistake"

    b. Wahrscheinlich hat ein Kind geweint
        probably   has a   child wept
        "Probably, a child wept"

    c. Früh am   Morgen hat ein   Eisbär   einen Mann gefressen
        Early in the morning has a.NOM   polar bear a.ACC man   eaten
        "Early in the moring, a polar bear has eaten a man"

As Fanselow (2002) and G. Müller (2003) observe, nominative subjects, dative arguments of unaccusative and passive predicates, and temporal and sentential adverbs have in common that they can be the structurally highest phrases in IP (even if they occupy different positions). Relative to (11), (9, 10) are easy to analyze: when C has no semantic or pragmatic feature, its EPP feature attracts the closest (=highest) category in IP. This is what one would expect, given the Minimal Link Condition (12): C cannot attract γ to its specifier position α in (13) if there is a β closer to C that can move as well.

(11) a. dass ein   Kind seinen Schlüssel verloren hat
         that a.NOM child his.ACC key   lost   has
         "that a child has lost his key"

      b. dass einem Studenten ein   Fehler aufgefallen war
         that a.DAT   student   a.NOM   mistake struck   was
         "that a student noticed a mistake"

      c. dass wahrscheinlich ein Kind geweint hat
         that probably   a   child wept   has
         "that probably a child has wept"

      d. dass früh   am Morgen   ein Eisbär   einen Mann gefressen hat
         that early in the morning a   polar bear a   man eaten   has
         "that a polar bear ate a man early in the morning"

(12)   MLC: A cannot attract B if there is a C, C closer to A than B, such that C can be attracted by A

(13)   [CP α COMP [IP β .... γ ...]]

This analysis of (9)-(10) in terms of an EPP-feature attracting the closest element is reminiscent of the analysis Holmberg (2000) proposes for Stylistic Fronting in Icelandic, in which an EPP-feature of I attracts the closest DP, PP,

adverb, or participle. In this sense, (9) and (10) exemplify Stylistic Fronting at the CP-level.

In the spirit of a proposal Bhatt (1999) made for Kashmiri, the analysis of type A and type B sentences can be unified: Σ can attract a category only if attraction establishes a checking relation, which presupposes that the features of the attractor and the attractee match. Therefore, C attracts the absolutely closest phrase β if C has no further features (=Type B, *Stylistic Fronting*). However, when C has an additional operator feature [+g], a checking relation can be established with those categories only that possess [+g] as well. Thus, β can be skipped in (14) if γ is the closest category with a matching specification of [+g]. Wh-phrases, and focal and topical material may cross higher phrases when C bears operator features (type A clauses, operator movement).

(14) $[_{CP} \alpha \, COMP[+g] \, [_{IP} \beta \, ... \, \gamma[+g] \, ... \,]]$

The idea that the highest element in IP moves to Spec,CP when C has an EPP-feature but no operator feature implies further predictions about what can show up in Spec,CP. Recall that arguments can be placed into a pre-subject position in German by scrambling (see Fanselow 2001, 2003a, Haider & Rosengren 1998, Grewendorf & Sabel 1994, Müller & Sternefeld 1993), e.g., in order to satisfy the word order principle that animate XPs precede inanimate ones (Hoberg 1981, G. Müller 2000). Animate objects can thus precede an inanimate subject as in (15a,c) without having any particular pragmatic function of their own, and they can be placed subsequently into Spec,CP in main clauses when the only attracting feature of Comp is the EPP-feature. There are no restrictions on the category and grammatical function of an element moved to Spec,CP by Stylistic Fronting—it merely must happen to be the highest element in IP.

(15) a. dass fast   niemandem   das   Medikament   geholfen   hat
        that nearly nobody.DAT the   medicine   helped   has
        "that the medicine nearly hasn't helped anybody"

     b. Fast niemandem hat das Medikament geholfen

     c. dass niemanden   der   Vortrag   geärgert   hat
        that nobody.ACC the.NOM talk   annoyed   has
        "that the talk hasn't annoyed anybody"

     d. Niemanden hat der Vortrag geärgert

Frey (2004) shows that topic phrases must c-command sentence level adverbs. When *mein Buch* "my book" is a topic as in (16), it must c-command *wahrscheinlich* (16a). (16b) and (16c) are not pragmatically felicituous continuations of the first sentence in (16). By being moved above the sentential adverb, the topic is the <u>highest</u> element in IP, so it will be placed into Spec,CP in main clauses such as (16d) even when C possesses an EPP feature only.

(16) Erzähl mir was über Dein Buch "Tell me something about your book"

     a. Ich denke, dass mein Buch vielleicht   ein   Litauer   publiziert
        I   think   that my   book perhaps   a   Lithuanian publishes
        "I think that a Lithuanian will perhaps publish my book"

     b. #Ich denke, dass vielleicht mein Buch ein Litauer publiziert

     c. #Ich denke, dass vielleicht ein Litauer mein Buch publiziert

     d. Mein Buch wird vielleicht ein Litauer publizieren

According to Frey (2004), topics occupy the highest position below C. They will thus be attracted to Spec,CP by a bare EPP-feature. Thus, there are also no pragmatic restrictions on what will be placed to Spec,CP by Stylistic Fronting.

3 PPT Movement in German Wide Focus Constructions

3.1 First facts

There is one class of German main clauses that is not covered by the model sketched in the preceding section. Surprisingly, even (2), repeated here as (17), turns out to be problematic in contexts when (2) felicitously answers questions such as *what did you do last weekend?* or *what happened last weekend?*, so that *ein Buch* is not the narrow focus of the utterance (and could be attracted relative to a [+foc] feature of C).

(17)   Ein BUCH hab   ich gelesen
        a   book   have I   read
        "I have read a book"

When the complete VP or IP is in focus, (18a) can be used, since focus movement to Spec,CP is optional in German. The direct object receives primary stress (e.g., by the NSR) since it is the most deeply embedded category in VP. A VP in focus can also be attracted to Spec,CP as in (18b), if C bears a [+foc] feature.

(18) a.   ich hab [$_{VP}$ ein BUCH gelesen]

     b.   [$_{VP}$ ein BUCH gelesen] hab ich

That (17) is a further option when VP/IP is in focus has not gone unnoticed (see, e.g., Büring 1996: 39). That a direct object can move to Spec,CP *at all* under such circumstances is surprising: the object DP is *not* the focus of the utterance (it is *part of the focus*). It should not have a [+foc] feature, so it is unclear how it can be attracted by C bearing [+foc]. (17) thus illustrates the "opposite" of pied-piping, *pars pro toto movement* (ppt-movement): XP seems to be attracted by a feature of a head, but only *part* of XP actually moves.

Direct objects can, in general, undergo ppt-movement. Both (19a) and (19b) allow wide focus. Only elements bearing primary stress undergo ppt-movement. Therefore, (19c) has no ppt-movement/wide focus interpretation: the PP bears narrow focus. This stems from the fact that the stress placement rules put the primary accent on the direct object rather than the prepositional object in a wide focus interpretation of (19a)[2].

(19) a.   Ich   hab   die Bücher   ins   Regal   gestellt
            I   have   the books   into-the shelf   placed
            "I put the books on the shelves"

      b.   Die BÜCHER hab ich ins Regal gestellt

      c.   #Ins Regal hab ich die Bücher gestellt

However, ppt-movement is not confined to direct objects. When the direct object is deaccented because it represents given information, primary stress is shifted, e.g., to the indirect object, which can then be preposed by ppt-movement: (20b) can answer (20a), i.e., the whole VP/IP (except for the object pronoun) is in focus. (Of course, (20c) is appropriate as well).

(20) a.   Was ist mit dem Buch passiert "What happened to the book?"

      b.   Meiner FREUNDIN hab   ich 's   geschenkt
            my.dat girlfriend   have I   it   given
            "I gave it to my girlfriend as a present"

      c.   Meiner Freundin geschenkt hab ich's.

Arguably, subjects may undergo ppt-movement as well if the objects are de-accented. (21b) can continue (21a) in a felicitous way. The primary accent on

---

[2] In contrast to what one would expect under a simplistic interpretation of the NSR.

Antje thus does not presuppose narrow focus on the subject, rather, the whole IP can be read as a comment on the sweater. Since the objects are deaccented, primary stress shifts to the subject, which ppt-movement will then place into Spec,CP. Given that the subject is also the highest element in IP, it is difficult to keep the effects of ppt-movement apart from those of "stylistic fronting".

(21) a.   Das ist aber ein schöner Pulli! "That is a really nice sweater"

    b.   ANTJE hat mir den geschenkt
        Antje   has me   it   given
        "Antje gave it to me as a present"

Even verbs can undergo ppt-movement if everything else is de-accented:

(22) a.   Was ist letzten Sonntag passiert? "What has happened last Sunday?"

    b   VERLETZT hab ich   mich
        hurt   have I   myself
        "I hurt myself"

3.2 A simple analysis

In ppt-movement constructions, part of an operator rather than the operator itself moves to Spec,CP when C possesses an operator feature. In a certain sense, ppt-movement data are thus reminiscent of *wh*-movement patterns such as (23)-(24). *Wh*-movement normally involves the displacement of a full *wh*-DP, but some *wh*-determiners can also move alone.

(23) a.   **Was für Bücher** hast du   gelesen?
        what for books   have you read

    b.   **Was** hast du **für Bücher** gelesen?
        "What kind of books have you read?"

(24) a. **Wieviel   Geld**   hast   du   dabei
         how much   money have   you   therewith

   b. **wieviel** hast du **Geld** dabei   [ok in certain dialects only]
      "How much money do you have with you?"

In minimalist theories of movement (Chomsky 1995), heads attract sets of formal features. Movement is covert if nothing else is displaced. Overt movement pied-pipes the phonological (and semantic) information linked to the attracted set of formal features. In the default case, the smallest unit with the attracted feature set (normally: a word) moves. Overt wh-attraction thus triggers the preposing of the *word* that bears the attracted feature (*was*, *wieviel*). as in (23-24b). In many cases, further principles require or allow that more material is pied-piped, yielding full phrasal movement as in the a-examples of (23, 24).

   In the optimal state of affairs, this concept of movement characterizes focus movement and topicalization as well. Overt topicalization and focus movement should therefore also either prepose the *word* bearing the topic (focus) feature, or some phrase dominating that word.

   While words bear the wh-feature as part of their lexical specification, focus and topic are marked prosodically in German. A word "bears" a focus feature (more precisely: a focus feature manifests itself on a certain word) if that word bears a particular pitch accent. The minimalist perspective implies that the word marking focus prosodically (=*Buch* in (25)) is the smallest unit that can move when a focus feature is attracted, unless the pied piping of larger categories is grammatically required. In (25) and the examples in 3.1., the smallest maximal projection dominating the word marking focus prosodically has to be pied-piped, because normally, maximal projections only move to Spec,CP in German.

(25)   a. [Ein BUCH]   hab   ich gelesen
          a   book   have I   read

b.   [ein BUCH gelesen] hab ich

A pitch accent on *Buch* can mark a (narrow) focus of the DP object, or focus of the VP and (wide) focus of the IP. The preposing of the object-DP as in (25a) can occur in *all* these focus constellations. The attraction of a word with a *formal* property (a pitch accent signalling focus) is what is relevant, and not the *semantic or pragmatic* status of the phrase that is fronted. The attracting [+foc]-feature of C is thus not an operator feature in a natural sense. It is a feature checking for the *marking* of operator status. Non-stylistic movement to Spec,CP is not operator movement (as we had suspected in sect. 2)—the features that C attracts belong to the formal rather than semantic-pragmatic aspect of language[3].

Given that focus marking is prosodic in German, the element that is attracted on the basis of a foc-feature is identifiable only *after* the computation of pitch accents. This presupposes the cyclic view of the interaction of syntax and phonology proposed by Bresnan (1972). In order to account for the pragmatic potential of (25a), the computation of the focus which is marked by the primary accent must also precede the (potential) movement of the focus marked element to Spec,CP. From a purely syntactic perspective, the choice among several concrete models seems to be of little importance. We can, e.g., assume a bottom-up computation of focus, which begins with the word bearing accent, and then works its way up the tree with rules for projecting focus marking on dominating categories, which depend on the structural position of the focus marked category, the deaccentuation-status of sister categories, etc.

---

[3]   The term ppt-movement introduced in Fanselow (2003b) thus refers to the relation between the formal operation and the semantic-pragmatic function of the clause only. We suspect that wh-movement is also due to the attraction of a wh-marker rather than to the need to move an operator, but a discussion of this issue is beyond the scope of the present paper. This view is reminiscent of Bayer (1996) and other approaches that claim that wh-movement serves the purpose of clausal typing (rather than scope assignment to operators).

As we have already said, such a theory of attraction must be complemented by a theory of pied-piping. Typically, the complete minimal maximal projection dominating the attracted word must be displaced. It is unclear, however, what the upper limits of pied-piping are. Thus, (25b) with a fronted VP is not only compatible with VP- and IP-focus. It can also answer a question such as *what have you read*, i.e., the fronting of a VP can occur in situations in which the object DP has a narrow focus.

3.3 Focus particles

The ppt movement idea helps to solve a number of further riddles of German syntax some of which we discuss in this subsection, and others, in section 5.

Consider (26) in this respect. It is ambiguous between the two interpretations (27a) and (27b), as noted in, e.g., Fanselow (1993). (27a) (="the only thing he never read is the bible") is unproblematic in a straightforward way: *nur die Bibel* constitutes the narrow focus of the sentence, and such a narrow focus can be moved to Spec,CP in all approaches. In the interpretation (27b), however, the scope of the focus particle comprises the *whole* verb phrase: the sentence is felicitous in a situation in which various religious activities are discussed, and in which it is claimed that one of these (reading in the bible) has never been carried out by the subject of the sentence.

(26)   Nur die Bibel hat er nie   gelesen
         only the bible has he never read

(27)   a. Only for x, x = the bible: he has never read x

         b. Only P, P = bible reading: he has never done P

In our approach, both readings are unproblematic: the different focus assign-ments share the location of the focus marker, so it is no wonder that attraction to Spec,CP will treat them in the same way[4].

Büring & Hartmann (2001) deny that *nur* and the object DP form a constituent in (26). In their account, *nur* is adjoined to the entire CP (with *die Bibel* appearing in Spec,CP) because they assume that focus particles cannot adjoin to arguments. If they are correct, (26) constitutes no additional argument in favor of ppt-movement, but stills falls in line with what we have said in the preceding subsections.

As Büring and Hartmann concede, the need to assume that V shows up in *third* position when focus particles are adjoined to CP is an expensive price to pay for restricting *nur*-adjunction to extended verbal projections. They claim that the adjunction of *nur* to the entire CP is motivated because there is no scope-reconstruction of clause-initial *nur* into the main clause. This claim is factually incorrect, however: (28) and (29) allow a reading in which the subject quantifier takes scope over *nur* + DP. This is hardly expected if *nur* +DP has not been moved to clause-initial position. Therefore, we prefer not to assume the CP-adjunction theory of (26). See also St. Müller (2005a) for more observations that show that *nur* adjoins to argument-DPs and PPs

(28)   Nur zu Weihnachten geht jeder dritte   in die Kirche
        only at Christmas   goes each third   in the church
        "For every third person x: x goes to church only at P, P = Christmas"
        "Only at P, P = Christmas: every third person goes to church at P"

---

[4] We need to assume that focus particles such as *nur* 'only' may attach to the focus marking category independent of their final scope, so that *nur* can affect VP despite the fact that it is attached to DP. Scope extension for focus particles is needed independent of the German examples under discussion, however, since Japanese focus postpositions such as *mo* may also attach to the direct object when they take scope over VP (Shin Ishihara, p.c.).

(29)   Nur   die Bibel   liest   kein frommer Christ
        only the bible   reads no   religious christ
        "Only for x = x the bible: no christ reads x"
        "There is no y, y a christ: y reads only x, x the bible"[5]

Scope reconstruction is, of course, possible with ppt-fronting as well, as (30)
shows with the interpretation indicated below the example.

(30)   Was seinen Hochzeitstag betrifft:/ "As for his wedding anniversary:"
        [Nur einen Blumenstrauß]   überreicht jeder dritte Ehemann
         only a   bunch of flowers hands over every third husband
        "For every third x, x a husband: x does only P, P = hand over a bunch of
        flowers"

4   Other Languages

In German main clauses, C may possess features by which it attracts focus
marked phrases. A phrase is focus marked if it bears the relevant pitch accent.
The word bearing this accent is the smallest unit that can be attracted. German is
not the only language with these properties. In their detailed analysis of Czech
focus placement, Lenertova & Junghanns (2004) observe that the focus
exponent may be moved to clause-initial position in all-focus utterances:

(31) [A: What's new? B:] (= their 25)
      a.   MArtu   jsem   potkala.
           Martha.ACC   aux1SG   met.SGFM

      b.   Potkala   jsem MArtu.
           "I met Marta"

Junghanns (p.c.) may be correct in pointing out that ppt-movement is more
widespread/less restricted in Czech than in German, but at the same time,

---

[5]   The two readings come with clearly different intonations, a fact we will not try to explain
      here.

Lenertova & Junghanns (2004) observe that the deaccentuation of the subject is often necessary for the availability of a ppt-interpretation in Czech as well, see (32) (=their (37)), similar to what holds for German.

(32) A: Co je nového? 'What's new?'   B:

    a.   **GUláš**   jsem   uvařila.
           goulash   aux.1SG   cooked
           "I cooked goulash"

    b.   #**GUláš** matka uvařila.   (#=pragmatically inappropriate)
           goulash mother cooked
           "Mother cooked goulash"

    c.   **Ma**tka uvařila   **GUláš**.

In Russian, objects can also be fronted in all focus utterances, as (33) illustrates. However, ppt-movement seems to go along with special pragmatic effects: (33) seems to express that answering the question is somewhat superfluous (because the answer is obvious, or irrelevant, Katja Jasinskaja, p.c.). The same seems to hold for Croatian (Damir Cavar, p.c.) and Polish (Joanna Błaszczak, p.c.). Hungarian appears to allow for ppt-movement without this additional pragmatic flavor (Beata Gyuris, p.c.).

(33)   Chto   delaet Petja? Gazetu   on   chitaet.
          what   made   Peter   Newspaper he   read

PPT-movement effects can perhaps also be found in Tangale. Kenstowicz (1985) and Tuller (1992) observe that phonological processes such as vowel elision and left line delinking apply in the verb phrase between the verb and the object, but these processes are blocked when the object is focused. According to Kenstowicz and Tuller, this blocking constitutes indirect evidence for movement of narrow focus objects. Hartmann & Zimmermann (2004) show that the

relevant phonological processes are also blocked when the whole verb phrase is focused. In our model (though not the one proposed by Hartmann and Zimmermann), this can be counted as an instance of ppt-movement.

5   More Constructions

5.1 Idioms

The behavior of idiomatic expressions may be particularly helpful in identifying the nature of syntactic movement processes. In general, only the whole idiom has a meaning, its parts lack an independent interpretation[6]. The displacement of parts of an idiom should therefore be possible only if attraction affects purely formal properties, and not when operator features in a strict sense are involved.

      Subjects (of intransitive verbs) that are part of an idiom go to Spec,CP easily, since they are attracted by the EPP-feature of Comp, as an instance of the "Stylistic Fronting" aspect of German verb second constructions:

(34) a.   das Ende der   Fahnenstange ist erreicht
             the   end   the.GEN   flag pole   is   reached
             "it's enough!"

       b. die Kuh ist vom   Eis
          the cow is off-the   ice
          "the problem is solved"

As G. Müller (2000a) and St. Müller (2002, 2003), among others, have observed, non-subject parts of idioms can be fronted as well. This is illustrated

---

[6] This view needs to be qualified, see Nunberg, Sag, & Wasow (1994) for an illustration and explanation of the fact that many idioms are semi-transparent. To the extent they have a semicompositional meaning, parts of idioms might play different information structure roles. The argument in the above section is therefore strongest for those idioms that defy a semi-decomposition such as *den Garaus machen* "to kill". I am grateful to Stefan Müller for pointing this out to me.

by the examples in (35). The original observation is attributed to Marga Reis by Büring (1996).

(35) a. schöne AUGEN hat er   ihr gemacht
         beautiful eyes   has he   her made
         "he made eyes hat her"

     b. den GARAUS hat er ihr gemacht
         the *garaus*   has he her made
         "he killed her"

The direct objects of (35) are not the highest IP-elements, so their fronting cannot be explained as an instance of "Stylistic Fronting". They can be attracted easily in the context of ppt-movement, however: the pitch accent on the object marks wide focus on the VP/IP. Thus, (35a) can, e.g., answer a question like *Why do you think he loves her?* The whole predicate is in focus, not the object DP (this would make little sense, since the object is just *part* of the idiom). If C attracts a formal feature (as we assume), the explanation of (35) is easy. If C would attract a focus *operator*, however, the analysis of (35) would be unclear, since the idiom parts are not meaning-bearing elements.

Not unexpectedly, idiom parts can be fronted in Czech as well, see (36) taken from Lenertova & Junghanns (2004) (=their 32). Similarly, the idiomatic reading does not get lost when part of the idiom is fronted in Russian, as in example (37) (Katja Jassinskaja, p.c.)

(36) A: Proč ses s ním tak pohádal?
      'Why did you have such a quarrel with him?' B:

      a. **BOUdu**   na mě ušil!
          hut..ACC   for me he-stitched

      b. Ušil na mě **BOUdu**!
          "He has cheated me!"

(37)   A   chto delali bjurokraty?
        And what did   burocrats
        Da-a... Palki oni   emu v   kolesa   vstavljali
        Well... Stick   they him in wheels inserted
        "Well, they put obstacles in his path"

Dutch allows the fronting of object parts of idioms as well, if the object bears the primary accent of the clause, as was already observed by Koster (1978)[7]:

(38)   Een poets (die)   heeft Jan me   nog nooit gebakken
        a   prank that   has   Jan me   yet never played
        "A trick, John has never played one upon me"

When one considers the partial fronting of idioms, it also becomes very clear that it is not just all focus or VP focus utterances in which movement to Spec,CP displaces elements bearing certain accents rather than meaning bearing units. It seems to hold generally that the word bearing the prosodic marker of some information structure function of XP can be fronted independent of the rest of XP. E.g., in the examples in (39), the idiomatic predicates may be contrastive topics (negation being the focus). Their accented elements can be fronted alone. Again, attraction affects the marker of contrastive topicality rather than the phrase that is the contrastive topic.

(39) a.   Ins Bockshorn jagen (intimidate, lit.: "into-the goat horn chase")
           Ins   Bockshorn hat er sich   nicht   jagen lassen
           into.the goat horn has he refl   not   chase let
           "He did not let himself be intimidated"

---

[7]   Shin Ishihara points out that Japanese appears to be an exception, as Miyagawa (1997) claims that the idiomatic meaning is lost when its part is long-distance (i.e., A'-)scrambled. With the clause-internal (i.e., A-)scrambling the idiomatic meaning is maintained.

b.   Am Hungertuch nagen (to be poor, lit: "to nag at the hunger-cloth")
      am   Hungertuch   müssen wir noch nicht nagen
      at.the hunger-cloth must   we   not   yet   nag
      "We are not yet really poor"

Of course, it is only the accented part of the predicate that can undergo ppt-movement. As was observed, e.g., by St. Müller (2002, 2003), the verb *cannot* be fronted to Spec,CP in such constructions: the examples in (40) are ungrammatical (or allow an irrelevant literal interpretation only):

(40) a. * Jagen hat er sich nicht ins Bockshorn lassen (cf. (39a))

     b. * Gemacht hat er ihr schöne Augen (cf. (35a))

     c. * Nagen müssen wir noch nicht am Hungertuch (cf. 39b))

     d. * Gemacht hat er ihr den Garaus (cf. 35b))

Multipart idioms shed an interesting light on the functioning of ppt-movement. Since semantic/pragmatic differences among the parts of the idiom can play a minor role only (in case of semi-transparent idioms) or no role at all, the mechanisms of fronting by formal features can be observed in its purest form. Consider now the following examples (@ stands for: "literal reading only")

(41) a.   Wir haben ihm den roten HAHN aufs   Dach   gesetzt
           we   have him   the   red   cock   on-the roof   put
           "We set his house on fire"

     b.   den roten Hahn haben wir ihm aufs Dach gesetzt

     c. @ aufs Dach haben wir ihm den roten Hahn gesetzt

(42) a.　Ich will dir keine STEINE in den Weg legen
　　　　　I want you no 　　stones 　in the way put
　　　　　"I don't want to place obstacles in your path"

　　　b.　Steine will ich dir keine in den Weg legen

　　　c. @ In den Weg will ich dir keine Steine legen

(43) a.　Er ist vom 　　Regen in die Traufe gekommen
　　　　　he is from.the rain 　in the eaves come
　　　　　"He jumped out of the frying pan into the fire"

　　　b.　Vom Regen ist er in die Traufe gekommen ...

　　　c. @ In die Traufe ist er vom Regen gekommen

(44) a.　Er sollte 　die Flinte nicht so schnell ins 　Korn werfen
　　　　　he should the gun 　not 　so fast 　　in.the corn throw
　　　　　"He should not give up so quickly"

　　　b.　Die Flinte sollte er nicht so schnell ins Korn werfen

　　　c. @ Ins Korn sollte er die Flinte nicht so schnell werfen

In spite of the fact that the prepositional object seems to have the more pro-
minent accent in (43)-(44) and other examples[8], the rule is that only the leftmost
accent bearing part of the idiom can be fronted in ppt-movement constructions,
see also St. Müller (2003) for this observation. In other words, when a certain
focused constituent such as *vom Regen in die Traufe kommen* contains two
accents, only the higher one of the two can be moved to the specifier of a C-
node that attracts a fm-feature. This is in line with what we expect, because
[+foc]-attraction must also be subject to the Minimal Link Condition.

---

[8] I am grateful to Susanne Trissler for first pointing this out to me.

Idiomatic expressions thus have helped to establish two points. Quite in general, C attracts words bearing the formal marking of some pragmatic function rather than the XP that bears the pragmatic function. When there is a choice, the leftmost/highest element bearing a relevant accent is attracted.

These finding also suggests a possible explanation for the observation made in Lenertova & Junghanns (2004) that ppt-movement of objects is best when the subject is a deaccented (or inaudible) pronoun. Perhaps, *all* accents in an all-focus utterance are visible to an attracting C, so that only the highest one can move because of the MLC. Therefore, objects cannot cross accented subjects (32). If the subject is deaccented, it will not block the movement of objects. Probably, subjects bearing an accent marking a pragmatic function different from the one that C attracts also trigger no intervention effects.

(45) (brought to my attention by Uwe Junghanns) might argue against this explanation. That (45) can be an all-new utterance is suggested by the fact that it could function as the first sentence of a newspaper article.

(45)   Eine   furchtbare Entdeckung machten drei   Kinder   gestern   in der
        a.ACC   horrible   discovery   made   three children yesterday   in the
        Waldstadt
        forest city
        "Three children made a horrible discovery yesterday in the Waldstadt"

The object *eine furchtbare Entdeckung* has crossed a non-deaccented subject in (45). This could show that transitive subjects do not exert intervention effects for the attraction of pitch accent of the object, but one might also consider *eine furchtbare Entdeckung* the topic of the sentence, if topicality is understood in a quite extended sense of aboutness. Under this analysis, (45) would not involve ppt-fronting in an all focus context.

5.2 Particles

The movement of verb particles to clause initial position constitutes yet another instance of ppt-movement. Like other Germanic languages, German possesses particle verbs such as *mit+nehmen* "with-take". If the particle is stressed, it is stranded in verb second constructions, as exemplified in (46b). Bierwisch (1963) states that particles cannot move to Spec,CP, and this verdict has frequently been repeated (Haider 1990, Fanselow 1993, Olsen 1997, Eisenberg 1999), despite the fact that it is wrong: particles do undergo movement to Spec,CP, as (46c) illustrates. An illuminating discussion of particle fronting, based on corpus research, can be found in S. Müller (2002, 2003).

(46) a.　dass　er　Antje　nicht　mitnimmt
　　　　　　that　he　Antje　not　　with-takes

　　　b.　Er　nimmt　Antje　nicht　mit

　　　c.　Mit　hat　er　sie　schon　genommen
　　　　　　with　has　he　her　well　　taken
　　　　　　"he took her with him"

The grammatical status of the verbal particle is controversial (see Haiden 2004 for an overview), but researchers seem to more and more concur in the view that the verb-particle combination is syntactic rather than lexical (see St. Müller 2002, 2003, Wurmbrand 2000, Zeller 1999). Independent of whether *mit* in *minehmen* is dominated by a word level verb projection or not, the particle is *not* the highest category in IP. Therefore, (46c) cannot exemplify the "Stylistic Fronting" type of German main clauses.

　　　In some cases, the fronted (semi-) particle has a clear semantic content of its own, which may be in focus or constitute a (contrastive) topic, as in (47).

(47)   ZuRÜCK   werde ich dich nicht bringen
        back   will   I   you not   bring
        "I won't take you back!"

However, most particle-verb-combinations have an idiomatic non-compositional reading only, in which the particle makes no identifiable semantic contribution. Nevertheless, particles of these verbs may be fronted, as Zeller (1999) and St. Müller (2002, 2003) have amply demonstrated, see also (48).

(48) a.   vor-haben (intend, lit.: "before-have")
           Vor   haben wir das schon gehabt
           before have   we   that well   had
           "We had intended that"

     b.   vor-machen (to fool, lit.: "before-make")
           Vor   kannst du   der wirklich nichts   machen
           before can   you her really   nothing make
           "You cannot really fool her"

     c.   an-kommen (be received, lit: "at-come")
           Gut an kommt dagegen   die Rede   von Hans
           well at   comes   in contrast the speech of Hans
           "The speech of Hans was well received, however"

The examples in (48) come with a clear information structure, in which the whole predicate (rather than the particle extracted from it) is in focus or constitutes a contrastive topic. Thus, *das vorhaben* (intending it) is the contrastive topic of (48a), while *vormachen* (fooling) is the topic of (48b). The same interpretation is possible for particles extracted from (partially) compositional combinations: a natural continuation of (47c) would be (49), which shows that *mitnehmen* "take along" rather than *mit* "with" is the contrastive topic of (46c).

(49) Aber getanzt hat er nicht mit ihr
      but   danced   has he   not   with her
      "But he did not dance with her"

The particle is not the head of the verb-particle combination, but it is the element that manifests the tone/an accent assigned to it. Particle preposing is thus a straightforward instance of ppt-movement.

5.3 Complex pragmatics in non-idiomatic constructions

The model developed here also helps to keep the definition of topic tidy. Consider the following dialogues in (50). The context questions make the predicates "not having recognized anyone" and "not having said anything" the topics of the following uttterances. As expected, the full topical VP can be moved to Spec,CP (the a"- and b"-examples), but the object can also go there alone. Under our account of attraction to Spec,CP, this is expected: fronting is licensed for all categories that dominate the element bearing the pitch accent marking the pragmatic function. If movement to Spec,CP would be operator movement, we would have to assume that negatively quantified DPs can be topics because of (50a', b'), clearly an unwelcome result.

(50) a. Auf dem Klassentreffen haben viele einige der Schulfreunde nach 20 Jahren nicht mehr wiedererkannt. Hat denn jemand gar niemanden wiedererkannt?
         "At the class reunion, many did not recognize some of their schoolfriends after 20 years. Did somebody not recognize anyone?"

   a'. Gar niemanden   hat nur der   Hubert wiedererkannt
        really nobody.DAT has only the.NOM Hubert recognized

   a". Gar niemanden wiedererkannt hat nur der Hubert
        "Only Hubert did not recognize anyone"

    b.   A: An der Dativ-Diskussion gestern haben sich fast alle beteiligt.
          "Nearly all participated in the dative discussion yesterday"

        B: Und wer hat nichts gesagt?
          "And who did not say anything?"

    b': A: Nichts   hat   eigentlich nur   die Sabine gesagt
           Nothing has   ptc   only the Sabine said

    b". A: Nichts gesagt hat eigentlich nur die Sabine
          "Well, only Sabine did not say anything"

PPT-topicalization is able to affect single words only:

(51) a.   War er Anarchist? "Was he an anarchist?"

     a'.   Häuser hat er jedenfalls   nie   angezündet
         houses has he in any event never   set on fire
         "He has never set houses on fire"

     b.   Ist er gebildet? "Is he educated?"

     b'.   Bücher hat er jedenfalls   viele gelesen
         books   has he in any event many read
         "At least, he has read many books"

     c.   Ist er ordentlich angezogen? "Is he dressed properly?"

     c'.   Krawatte trägt er jedenfalls   wieder mal   keine
         tie   wears he in any event again   once   no
         "Again, he does not wear a tie"

Discontinuous noun phrases as we find them in (51) are common in German (see Fanselow 1988, Riemsdijk 1989, Fanselow & Cavar 2002, van Hoof 2004). Typically, the left part of the discontinuous DP is a contrastive topic, whereas the right part is a narrow focus, as in (52).

(52)   Drosseln hab   ich zwei gesehen, Schwalben aber   fünf
        thrushes have I   two   seen   swallows but   five
        "I have seen two thrushes, but five swallows"

The examples in (51) share this basic alignment of the contrastive topic part of the utterance at its left periphery, and the focal part at its right edge, but in the contexts given, the contrastive topic is not constituted by the preposed noun, but rather by the predicate dominating it before movement (setting houses on fire, wearing ties, etc.). Contrastive topicalization of a VP can be marked by only preposing the head of the direct object. This is a clear example of ppt-movement that becomes particularly obvious when one considers sentence pairs with explicit contrasting of topics such as (53)

(53)   Ordentlich gekämmt war er bestimmt, aber Krawatte hat er wieder
        Properly   combed was he certainly but tie   has he again
        mal   keine getragen
        once no   wore
        "Though his hair was certainly combed properly, he again wore no tie"

Together with the preposing of particles, the contrastive fronting of a noun constitutes the purest examples of ppt-movement, then: only the morpheme that bears the relevant accent undergoes fronting.

6   Remnant Movement?

The preceding sections have argued that the preposing of focus and topic phrases does not come about by the attraction of an operator feature– rather, the category that marks the pragmatic function prosodically is attracted. It may (and sometimes has to) pied-pipe larger categories, including the full phrase that is a focus/topic. Of course, one can envisage alternative descriptions of the data, and one alternative that deserves special attention is remnant movement.

Descriptively, ppt- and remnant movement have the same overt effect: a category X has a certain pragmatic value and should move to some position, but overt displacement affects part of the phonological material of C only. Remnant movement was proposed as a tool for German syntax by Thiersch (1985) and den Besten & Webelhuth (1987,1990). See G. Müller (1998) for a detailed discussion (and Fanselow 2002a for a critique). In a remnant movement construction, phrases can be extracted from VP before the latter category moves to Spec,CP, as shown in (54).

(54) a.   [$_{IP}$ er nicht[$_{VP}$ sie geküsst hat]]   →
              he not   her kissed   has

     b.   [$_{IP}$ er nicht sie$_i$ [$_{VP}$ t$_i$ geküsst hat]]   →

     c.   hat [$_{IP}$ er nicht sie$_i$ [$_{VP}$t$_i$ geküsst _]]   →

     d.   [$_{VP}$ t$_i$ geküsst] hat er sie$_i$ nicht   →
           "he has not kissed her"

Instead of assuming that the bearer of the pitch accent is attracted, a more classical remnant moement account seems possible, in which our standard example might have the structure in (55) (coming close to what was proposed in Fanselow 1993): the fronted material looks like a DP, but in fact, it contains an empty verbal head, making it a VP. Under this analysis, the element moved to Spec,CP *is* the focus phrase.

(55)   [$_{VP}$ [$_{DP}$ Zeitung] t$_V$] hab ich gelesen$_V$

The remnant movement account introduces the process of removing the verb from the VP as a factor influencing the acceptability of the construction. According to Wurmbrand (2001), (56a) is indeed better than (56b).

(56) a. Vor   machte er ihr nichts
          before made   he her nothing

   b. Vor   hat er ihr nichts   gemacht
       before   has he her nothing made
       "He has not fooled her"

In (56a), the underlying VP is something like [$_{VP}$ *ihr nichts vor-machte*], out of which *ihr* and *nichts* can be scrambled, while *machte* moves to C because of the verb second property. All processes involved are well-motivated in German syntax. This is different with (56b): here, *gemacht* is not the finite verb, hence it does not move to C. It can leave [$_{VP}$ *ihr nichts vor-gemacht*] only if we assume there is a rightward movement process for non-finite verbs, adjoining them to I, e.g., This process is not independently motivated, and this might account for any contrast between (56a) and (56b). However, we carried out a questionnaire study and failed to observe any difference in acceptability[9] between these structures.

The availability of a movement operation that extracts non-finite verbs from VP is however crucial for the remnant movement approach. Without such a movement, the phonological material of a DP could never constitute an VP. Haider (1993) and Koopman (1995) argue that there is no such movement of non-finite verbs in German and Dutch. Certain verbs come with two prefixes rather than one, and they typically must not appear in second position, as the contrast between (57a) and (57b-d) suggests. Haider and Koopman derive this and similar contrasts from the assumption that verbs like *voranmelden* cannot

---

[9] 48 Subjects (university students) rated 100 sentences on a 7 point scale (1 completely ungrammatical, 7 fully grammatical). Among these 100 sentence, there were 12 items related to the distinction in (56), 4 items belonging to the condition in which the lexical verb was clause final, and 4 items each relating to two conditions in which the lexical verb appeared in second position. Average acceptability of the sentence was between 4.6 and 4.8. The conditions did not differ from each other statistically.

undergo overt movement at all. If this is correct, the wellformedness of (57e) suggests that *voranmelden* has not moved overtly, that is, (57e) supports the idea that non-finite verbs do not adjoin to the heads selecting them in overt syntax.

(57) a. dass er sein Kind   vor-an-meldet
         that he his child   pre-at-reports
         "that he pre-registers his child"

   b. * er voranmeldet sein Kind

   c. * er meldet sein Kind voran

   d. * er anmeldet sein Kind vor

   e.   er wird sein Kind voranmelden können
         he will   his child pre-at-report can
         "he will be able to pre-register his child"

A discussion of the contrast in (58) can be found in Haider (1997) and Meinunger (2001). Apparently, main verbs must not move out of the scope of certain operators such as *mehr als* "more than", which implies that these verbs cannot enter simple tense main clauses (58b). (58a), on the other hand, is grammatical since finite verbs stay *in situ* in embedded clauses. By the same logic, (58c) implies that non-finite verbs do not undergo overt head movement either.

(58) a. dass er den Gewinn [mehr als   verdreifachte]
         that he the profit   more than tripled
         "that he more than tripled his profit"

   b. * er verdreifachte seinen Gewinn mehr als t

   c.   er hat seinen Gewinn mehr als   verdreifachen können
         he has his   profit   more than triple   could
         "he has been able to more than triple his profit"

Data such as (57)-(58) suggest that movement to C is the only type of overt V movement in German. Because German is an OV-language, this is in line with the typological generalizations uncovered by Julien (2002). If there is no overt movement of non-finite verbs, many ppt-data cannot be reanalyzed as involving remnant movement, because there is no process removing V from VP[10].

(59) constitutes a further problem for the remnant movement theory. If the preposed material in (59c) is to be analyzed as an instance of remnant VP movement, one not only has to assume that *gesetzt* may move out of VP– we also need to extract the PP *aufs Dach* from VP. There is no independently motivated movement transformation which could do this (note, e.g., that (59b) is indeed ungrammatical). In particular, scrambling cannot be invoked, since the PP is part of the idiom, and therefore meaningless. Scrambling never affects parts of idioms.

(59) a.   Wir haben ihm den roten HAHN aufs   Dach   gesetzt
          we   have him the red   cock   on-the   roof   put
          "we set his house on fire"

   b. ?*wir haben ihm aufs Dach [$_{VP}$ den roten Hahn t gesetzt]

   c.   Den roten Hahn haben wir ihm aufs Dach gesetzt

Consider also (60) in this respect. The verb phrase is merged as [$_{VP}$ [$_{DP}$ *keine Krawatte*] *getragen*]. The noun must then be separated from the determiner in the derivation of (60) in remnant movement theory. Since German disallows left branch extractions, the separation can only arise by extracting *Krawatte* from the

---

[10] Drawing firm conclusions from (57)-(58) and similar data is difficult, however. E.g., the constellation in (57) could also be explained if we assume that particle stranding is obligatory for V to C movement only, but optional otherwise. (57e) could then involve the non-stranding version of verb movement to I, while the optionally stranded particle could still be fronted in (57b) in the context of the movement of a radically evacuated VP.

DP, yielding [_{VP} *Krawatte* [_{VP} [_{DP} *keine* t] *getragen*]. The target remnant VP [_{VP} *Krawatte* [_{VP} __]] can then be generated by extracting the remnant DP and the verb from VP. This derivation of (60) leaves it open why neither [_{VP} [_{DP} *keine* t] *getragen*] nor [_{VP} [_{DP} *keine* t]] can be found in VP-contexts in German, as (61) illustrates.

(60)   Krawatte hat er keine getragen
        tie   has he no   wore
        "He hasn't worne a tie"

(61) * keine getragen hat er Krawatte
     * keine hat er Krawatte getragen

The attempt to reanalyze ppt-movement as radical remnant movement thus seems to fail because more often than not the movement operations required for evacuating, e.g., VP prior to movement to Spec,CP are not licensed.

    Our conclusion only holds for what G. Müller (2002) calls "primary" remnant movement, introduced by Thiersch (1985) and den Besten & Webelhuth (1987) as the interaction of independently motivated operations. "Secondary" remnant movement was proposed by Kayne (1998): it mainly serves to restore constituent order when theory-driven movement operations have yielded incorrect linearizations. Müller (2002) shows that primary and secondary remnant movement have quite different properties. Secondary remnant movement is never feature driven, and is quite unconstrained in grammatical terms.

    Obviously, one could postulate a set of secondary remnant movement operations in order to avoid ppt-movement. These operations would extract the verb and other deaccented material from VP, in order to create a VP which contains no phonetic material but a maximal projection with the primary accent. They would neither be motivated independently, not would they serve the need

of re-establishing constituent order (as in Kayne 1998). Rather, they only apply because of the need of creating, e.g., focused XPs that contain no overt material but the focus marked phrase. Since the theoretical properties of these secondary operations cannot be those of standard movement (as G. Müller 2002 has shown), the major problem of such an approach is obvious: it is hard to see how it could at all be refuted empirically. Its sole motivation would be to maintain a theoretical postulate (C attracts a focus or topic operator) in the light of empirical counterevidence[11].

7   Some Concluding Remarks

In section 2, we introduced two types of filling the first position in German main clauses: the attraction of the element closest to Spec,CP in case C has no operator feature, and the attraction of an f-operator in case C has an attracting f-feature. This paper has shown that we can maintain this basic dichotomy, but one needs to revise the idea that C attracts operators. Rather, C attracts the word/phrase that bears the marking of an operator.

There are at least two issues that deserve further attention. First, the ppt-movement approach implies that one and the same information structure constellation (say: focus on VP) can be expressed by several different movement operations: it would suffice to prepose the DP-object, but the full VP may be pied-piped, too. Is the choice among these constructions really optional? Note that the pertinent problem is not confined to the ppt-movement theory: in all accounts, the apparent *overall* optionality of focus movement constitutes a problem. Furthermore, in a model in which the driving force for movement is the presence of a focus or topic marker, there is no obvious pre-theoretic sense

---

[11]   The partial deletion approach of Fanselow & Cavar (2002) could also replace ppt-movement in principle. Just like secondary remnant movement, the partial deletion approach suffers from the fact that it is not restricted enough.

in which the fronting of the object should be more economical than the fronting of the full VP, or vice versa: mere object fronting makes VP discontinuous but leaves base order relations between the verb and the subject intact, while full VP fronting does not create a discontinuous VP, but changes the serialization of *all* elements in VP relative to the rest of the clause. In a way, then, ppt-movement theory predicts the range of choice we have for pied piping in the context of focus or topic movement. In contrast, if attraction would be based on semantic features, DP-fronting for marking VP-focus is difficult to understand and should be highly marked, in contrast to facts.

However, it may be true that (62b) is in a sense more "emphatic" than both (62a,c). The question then is how that impression can be made precise, and how it will formally figure in the attraction account.

(62) a.   ich hab   ein Buch gelesen
          I   have   a   book   read

b.   ein Buch hab ich gelesen

c.   ein Buch gelesen hab ich

Some of the examples we have discussed may be characterized by a "topic feature within focus construction" constellation (Caroline Féry, p.c., Lenertova & Junghanns 2004), that may also help to choose between the various fronting options, but, as we have said above, it is dubious whether that affects *all* instances of ppt-movement.

A second area for future research is multiple fronting. For example, as St. Müller (2003) observes, particle fronting co-occurs with a fundamental enigmna of German main clauses, viz. "multiply" filled Spec,CP positions, see (63) based on the complex verbs *vor-haben* "to plan" and *an-haben* "to wear".

(63) a. [heute abend] [einiges] [vor] dürfte Antje schon noch haben
        tonight   plenty PTC might Antje well   still   have
        "Antje may very well have a couple of nice plans for tonight"

    b. viel an hatte er ja   nicht mehr
        much on had he well no   longer
        "he did not really wear a lot"

In our model, the particles are fronted because they are attracted as bearers of the relevant pitch accent. It is unclear, though, what category has been pied-piped in the context of this attraction process. Fanselow (1993) and St. Müller (2003, 2005) offer arguments suggesting that the material preceding the finite verb in (63) is a VP lacking an overt verbal head. However, the arguments against a remnant movement analysis apply to (63) as well, so it is unclear how VP got decapitated in (63). St. Müller (2005) proposes a theory that makes use of some of the technicalities of HPSG that cannot be translated into minimalist and other movement based accounts.

References

Bayer, Josef. 1996. *Directionality and Logical Form*. Dordrecht: Kluwer.

den Besten, Hans. 1989. *Studies in West Germanic Syntax*. Amsterdam: Atlanta.

den Besten, Hans & Gert Webelhuth. 1987. Remnant topicalization and the constituent structure of VP in the Germanic SOV languages. Paper, presented at the X. GLOW Colloquium, Venice.

den Besten, Hans & Gert Webelhuth. 1990. Stranding. In *Scrambling and Barriers*, eds. Günter Grewendorf & Wolfgang Sternefeld. 77-92. Amsterdam: Benjamins.

Bhatt, Rakesh Mohan. 1999. *Verb Movement and the Syntax of Kashmiri*. Dordrecht: Kluwer.

Bierwisch, Manfred. 1963. *Grammatik des deutschen Verb*s (=Studia Grammatica II). Berlin: Akademie Verlag.

Bierwisch, Manfred 1968. Two Critical Problems in Accent Rules. *Journal of Linguistics* 4. 173- 178.

Bresnan, Joan. 1972. On Sentence Stress and Syntactic Transformations. In *Contributions to Generative Phonology*, Michael Brame (ed.), 73-107. Austin: University of Texas Press.

Büring, Daniel. 1996. *The 59th Street Bridge Accent. On the Meaning of Topic*. Dissertation, Universität Tübingen.

Büring, Daniel, & Katharina Hartmann. 2001. The Syntax and Semantics of Focus-Sensitive Particles in German. *Natural Language and Linguistic Theory* 19: 229-281.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20,* Ken Hale and Samuel Keyser (eds)*,* 1-58. Cambridge: MIT Press.

Chomsky, Noam. 1995. Categories and Transformations. In *The minimalist program,* 219-394. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by Ste*p, ed. Roger Martin, David Michaels & Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.

Eisenberg, Peter. 1999. *Grundriß der deutschen Grammatik*, volume 2. Der Satz. Stuttgart, Weimar: Verlag J. B. Metzler.

Fanselow, Gisbert. 1988. Aufspaltung von NP und das Problem der 'freien' Wortstellung. *Linguistische Berichte* 114: 91-113

Fanselow, Gisbert. 1993. Die Rückkehr der Basisgenerierer. *Groninger Arbeiten zur Germanistischen Linguistik* 36: 1-74.

Fanselow, Gisbert. 2001. Features, θ-roles, and free constituent order. *Linguistic Inquiry* 32: 405-437.

Fanselow, Gisbert. 2002. Quirky Subjects and Other Specifiers. In *More Than Words* eds. Barbara Kaufmann & Barbara Stiebels, 227-250. Berlin: Akademie Verlag.

Fanselow, Gisbert. 2002a. Against Remnant VP Movement. In *Dimensions of Movement,* eds. Artemis Alexiadou, Elena Anagnostopoulou, Sjeff Barbiers, & Hans-Martin Gärtner, 91-127. Amsterdam: Benjamins.

Fanselow, Gisbert. 2003. Münchhausen-Style Head Movement and the Analysis of Verb Second. *UCLA Working Papers in Linguistics*.

Fanselow, Gisbert. 2003a. Free Constituent Order. A Minimalist Interface Account. *Folia Linguistica*.

Fanselow, Gisbert. 2003b. Surprising Specifiers and Cyclic Spellout. In *Generative Linguistics in Poland* 5: 30-46.

Fanselow, Gisbert & Damir Ćavar. 2002. *Distributed Deletion*. In *Theoretical Approaches to Universals*, ed. Artemis Alexiadou, 65–107. Amsterdam, Benjamins.

Frey, Werner. 1993. *Syntaktische Bedingungen für die Interpretation*. Berlin: Akademie-Verlag.

Frey Werner. 2004. A Medial Topic Position for German, *Linguistische Berichte* forthcoming.

Gärtner, Hans.-Martin. & Markus. Steinbach. 2003. What do Reduced Pronominals Reveal about the Syntax of Dutch and German? Part 1: Clause-Internal Positions. *Linguistische Berichte* 195: 257-294. Part 2: Fronting. *Linguistische Berichte* 196: 459-490.

Grewendorf, Günther & Joachim Sabel. 1994. Long scrambling and incorporation. *Linguistic Inquiry* 25: 263-308.

Haider, Hubert. 1990. Topicalization and other Puzzles of German Syntax. In Günther Grewendorf and Wolfgang Sternefeld (eds.), *Scrambling and Barriers*, Linguistik Aktuell, 93–112, Amsterdam, Philadelphia: John Benjamins.

Haider, Hubert. 1993. *Deutsche Syntax - generativ*: Narr, Tübingen.

Haider, Hubert. 1997. Typological implications of a directionality constraint on projections. In *Studies on universal grammar and typological variation*, ed. Artemis Alexiadou & Tracy Hall, 17-33. Amsterdam: Benjamins.

Haider, Hubert & Inger Rosengren. 1998. Scrambling. *Sprache und Pragmatik* 49. Lund: Germanistisches Institut.

Haiden, Martin. 2004. Case 111: Verb Particle Constructions. *SynCom Project*.

Hartmann, Katharina & Malte Zimmermann. 2004. Focus Strategies in Chadic – the Case of Tangale Revisited. This volume.

Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. München: Hueber

Holmberg, Anders. 2000. Scandinavian Stylistic Fronting: How any category can become an expletive. *Linguistic Inquiry* 31: 445-483.

van Hoof, Hanneke. 2004. Split Topicalization (Case 101). *The SynCom project*.

Jacobs, Joachim. 2001. The dimensions of topic-comment. *Linguistics* 39: 641-682.

Julien, Marit. 2002. *Syntactic Heads and Word Formation*. Oxford: OUP.

Kayne, Richard. 1998. Overt vs. Covert Movement. *Syntax 1*.

Kenstowicz, Michael. 1985. The Phonology and Syntax of WH-Expressions in Tangale. In *Studies in the Linguistic Sciences* 15: 79-91

Koopman, Hilda. 1995. On verbs that fail to undergo Verb-Second. *Linguistic Inquiry* 26:137-163.

Koster, Jan 1978. *Locality Principles in Syntax*. Dordrecht: Foris.

Lenertova, Denisa & Uwe Junghanns. 2004. Left Peripheral Focus Exponents. Ms., Universität Leipzig.

Lenerz, Jürgen 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Narr.

Meinunger, André. 2001. Restrictions on verb raising. Ms., ZAS Berlin.

Müller, Gereon. 1998. *Incomplete category fronting*. Dordrecht: Kluwer.

Müller, Gereon. 2000. Optimality, Markedness, and Word Order in German. *Linguistics* 37:777-818.

Müller, Gereon. 2000a. Idioms and transformations. Paper, presented at the 2000 GGS meeting, Potsdam.

Müller, Gereon. 2002. Shape Conservation and Remnant Movement. In *Dimensions of Movement,* eds. Artemis Alexiadou, Elena Anagnostopoulou, Sjeff Barbiers, & Hans-Martin Gärtner. Amsterdam: Benjamins.

Müller, Gereon. 2003. Verb Second as vP First. *Journal of Comparative Germanic Grammar*.

Müller, Gereon & Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24: 461-507.

Müller, Stefan. 2002. *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German.* Studies in Constraint-Based Lexicalism, Nr. 13, Stanford: CSLI Publications.

Müller, Stefan. 2003. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31: 29-62.

Müller, Stefan. 2004. Zur Analyse scheinbarer V3 Sätze. Ms.

Müller, Stefan. 2005. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte*.

Nilsen, Oystein. 2003. *Eliminating Positions*. LOT, Utrecht.

Nunberg, Geoffrey, Sag, Ivan A. & Wasow, Thomas. 1994. Idioms. *Language* 70(3), 491–538.

Olsen, Susan. 1997. Zur Kategorie Verbpartikel. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 119, 1–32.

Riemsdijk, Henk van 1989. "Movement and regeneration". In: *Dialectal variation and the theory of grammar,* P. Benincà (ed.) Dordrecht: Foris, 105-136.

Rizzi, L. 1990. *Relativized Minimality*, MIT-Press, Cambridge, Mass.

Thiersch, Craig. 1978. *Topics in German Syntax*. Dissertation, MIT, Cambridge, Mass.

Thiersch, Craig. 1985. VP and Scrambling in the German Mittelfeld. Ms., University of Tilburg.

Travis, Lisa 1984. Parameters and Effects of Word Order Variation. Unpublished doctoral dissertation, MIT, Cambridge, MA.

Tuller, Laurice. 1992. The syntax of postverbal focus constructions in Chadic. *Natural Language and Linguistic Theory* 12: 303-334

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York, Oxford UP.

Wurmbrand, Susi. 2001. The Structure of Particle Verbs. Ms.

Zeller, Jochen. 1999. *Particle Verbs, Local Domains, and a Theory of Lexical Licensing*.Dissertation, University of Frankfurt, Frankfurt/Main.

Zwart, C. Jan-Wouter. 2001. Syntactic and Phonological Verb Movement. *Syntax* 4:36-62.

Gisbert Fanselow
Institut für Linguistik
Universität Potsdam
Postfach 601553
14415 Potsdam
Germany
fanselow@rz.uni-potsdam.de
http://www.ling.uni-potsdam.de/~fanselow

German Sentence Accent Revisited[*]

Caroline Féry and Laura Herbst

Universität Potsdam

Results of a production experiment on the placement of sentence accent in German are reported. The hypothesis that German fulfills some of the most widely accepted rules of accent assignment—predicting focus domain integration—was only partly confirmed. Adjacency between argument and verb induces a single accent on the argument, as recognized in the literature, but interruption of this sequence by a modifier often induces remodeling of the accent pattern with a single accent on the modifier. The verb is rarely stressed. All models based on linear alignment or adjacency between elements belonging to a single accent domain fail to account for this result. A cyclic analysis of prosodic domain formation is proposed in an optimality-theoretic framework that can explain the accent pattern.

Keywords: Prosody, Syntax, Information structure

1 Introduction

As predicted by most models of sentence accent placement, all-new German VPs consisting of an argument and a verb often have their main accent on the argument. Sentences in which a modifier is inserted between the argument and the verb are also accented on the modifier in addition to having a prenuclear accent on the argument. The verb is rarely accented. Selkirk's (1995) theory accounts for this pattern since the pitch accent on the argument is allowed to project on the whole VP, even if the modifier intervenes between argument and

verb, but Gussenhoven's Sentence Accent Assignment Rule (SAAR), or any other theory which claims that a focused (or new) adjunct can never project its accent—and most theories make this prediction—fail to explain this pattern. SAAR requires constituents to be adjacent in order to form a focus domain and all OT theories making use of Alignment constraints fall into this trap as well. A remedy could lie in a revision of the absolute prohibition of accent projection from an adjunct, or, alternatively, small focus domains (on the modifier) could be allowed to be embedded in larger ones (the VP) on a regular basis. To this aim, a cyclic account of prosodic domain formation is necessary. This paper first gives a review of past approaches to sentence accent (section 2). It then identifies the problem to be solved and presents the results of a production experiment (section 3). Finally, it gives a solution in terms of a cyclic OT analysis (section 4).

2   Background

2.1 Nuclear Stress Rule

From the 1960s on, German has played a prominent role in the discussion of sentence stress in a generative perspective, because, despite some similarities, stress location differs in crucial ways from English. The Nuclear Stress Rule (NSR), as formulated by Chomsky & Halle (1968), assigns main stress to the rightmost constituent (major-class word) in the sentence, on the basis of the surface linear ordering. This very simple principle accounts for most English utterances. In an optimality-theoretic model (OT, Prince & Smolensky 1993, McCarthy & Prince 1993a), the tendency can be formulated as a straightforward Alignment constraint (see McCarthy & Prince 1993b, Truckenbrodt 1999, and Samek-Lodovici 2004 for applications), requiring sentence stress to appear on the rightmost Prosodic Word, as in (1).

(1)   ALIGN-R (Intonation Phrase, main stress, Right)

Main stress is on the last Prosodic Word of the Intonation Phrase.

German has two classes of sentences, those fulfilling the NSR and those violating it.[1] Krifka (1984) lists minimal pairs, illustrating both behaviors of sentence stress placement.

(2) a. NSR is fulfilled   b. NSR is violated
       Lena liegt auf dem SOFA.   Lena hat auf dem SOFA gelegen.
       Lena lies on the   sofa   Lena has on the   sofa   laid

       Ede fährt jeden Tag drei STUNDEN. Ede ist nach FRANKFURT gefahren.
       Ede drives each day three hours   Ede is to Frankfurt driven

---

[1] According to Kiparsky (1966:81ff) the German syntactic constructions really divide into two groups. One group of syntactic constructions (called Nom, D and Sentence) gets final stress, and the other group (called VP and S) is initially-stressed. Kiparsky's examples are reproduced here.

Kiparsky (1966)

	2	1	2	1	2	1
Nom:	die dicke	Emma,	Karl der	Große,	der Mann	aus Rio
	the fat	Emma,	Charlemagne,		the man	from Rio

			2			2	1	
D:	dass (ein	Schüler	jede	Woche	einen	Aufsatz)	schreiben	muss
	that (a	pupil	every	week	a	report)	write	must

	2		1		2	1
Sentence:	Waldemar	spielt	Theater;	die	Katze lief	weg
	Waldemar	plays	theater;	the	cat ran	away

		1	2
VP:	Er hat	(schimpfen	wollen)
	He has	(to curse	wanted)

		1	2
S:	Er wird	(Purzelbäume	schlagen)
	He will	somersaults	beat

Es ist drei Gramm GOLD.   Es ist drei GRAMM schwer.
It is three grams gold   It is three grams heavy

Align-R (and the NSR) makes the right predictions for the expressions in (2a), but for the examples in (2b), this constraint fails. The more complex behavior of German sentence stress as compared to English has been met in different ways by different authors. Beside Kiparsky's solution (see footnote 2), Bierwisch (1968) has proposed to assign sentence stress in German at a non-surface level. Before transformations, stress is rightmost, after transformations, it is not necessarily so any longer (see also Fanselow, this volume, for a movement-based analysis of stress in German).

The assumption that stress is assigned at a deeper representation in the grammar has also been defended by Bresnan (1971, 1972) for English to account for examples like those listed in (3) to (6), first discussed by Newman (1946). In the (a) version, stress is non-final, but in the (b) sentences, it is final.

(3)   a.   George has PLANS to leave.
      b.   George has plans to LEAVE.

(4)   a.   Helen left DIRECTIONS for George to follow.
      b.   Helen left directions for George to FOLLOW.

(5)   a.   Whose UMBRELLA have I taken?
      b.   Whose umbrella have I TAKEN?

(6)   a.   I asked what BOOKS Helen had written.
      b.   I asked what books Helen had WRITTEN.

In Bresnan's account, stress assignment in (3a) to (6a) fulfills the NSR, but only at deep structure, in which the capitalized constituents are final. After stress assignment has applied, equi-deletion applies in (3a) and (4a), and wh-

movement in (5a) and (6a), delivering a surface structure in which stress is somewhere else than the final position.

Numerous counterexamples, both to the NSR and to Bresnan's amendment, have been discussed in the literature (see for instance Berman & Szamosi 1972 and Gussenhoven 1992). Gussenhoven (1992) shows that an example like (7b) cannot be accounted for in Bresnan's framework. Before wh-movement, NSR assigns stress to *countries*, but *coffee* bears the nuclear accent on the surface.[2]

(7) a. Coffee is grown in tropical countries.
     b. In which countries is coffee grown?

Even if the NSR applies strictly at the surface, the accent pattern is still unaccounted for. It is not *grown* which is mainly stressed, though it should be if one takes the fact that it is the final major class word into account. Gussenhoven shows that both Chomsky & Halle's and Bresnan's analyses are based on the wrong premises, and proposes that accent placement is assigned non-cyclically and at least in part according to the predicate-argument structure of the sentence.

2.2   Argument-stressing (integration)

Schmerling (1976) observes that the predicate-argument-structure plays a crucial role, both in English and in German. The fact that the NSR applies more successfully in English than in German is explained by the more frequent occurrence of an argument in the sentence-final position in English. Because of the verb-final pattern of the German embedded clauses and the placement of

---

[2] Gussenhoven considers not only the nuclear stress (the sentence final one), but also the prefinal ones. He differs in that from Bresnan and many other researchers who assume that there is only one main accent in a sentence. In all accounts, however, main sentence accent is assimilated to the nuclear stress, which is the last pitch accent.

nonfinite verbs in final position in all clauses, arguments of verbs are often non-final in German and, since they are the depository of sentence accent, NSR is violated. In the first pair of sentences in (2), violation of the NSR correlates with the placement of the verb in the final position and with the simultaneous locations of the arguments in the preverbal position. As a further example, consider the sentences in (8). The first sentence, a V2 sentence, has a final argument, but the second sentence has a final participle, and two pre-verbal arguments (or a complex one if the whole journey is considered as one argument). The preverbal argument bears main sentence stress.

(8) a. Mein Flugzeug hatte zwölf Stunden VERSPÄTUNG
         my   plane   had   twelve hours   delay

     b. Ich bin nämlich gestern   von Berlin nach BEIJING   geflogen
         I   am namely yesterday from Berlin to   Beijing   flown

Krifka (1984), von Stechow & Uhmann (1986) and Cinque (1993) among others have formulated different versions of a rule assigning stress to the argument in a German predicate-argument structure. In terms of constraints, the accenting of an argument can be formulated as in (9), see also Büring (2001).

(9)   Stress-Argument
       In a predicate-argument structure, stress lies on the argument.

If an argument cannot be stressed—because it is a function word, because it is part of the background, or because there is no argument in the sentence—the location of the main stress is decided by ALIGN-R. Rightmost stress is thus not eliminated from the analysis, but is a case of Emergence of the Unmarked: everything else being equal, stress is on the rightmost constituent.

Many researchers after Schmerling have claimed that English also relies on the predicate-argument structure to assign sentence stress, much in the same way as German does. But because of the different linearization of the constituents, the predictions of the NSR and those of STRESS-ARGUMENT only rarely differ. We have seen Bresnan's examples in (3) to (6). Returning to Gussenhoven's counterexamples in (7), *coffee*, as an argument of the passive verb *grown* gets the main stress.

Another context which reveals the stable preference for argument stressing in English has been discussed by Selkirk (1995). Like many other researchers, she assimilates the focused part of a sentence to the answer to a preceding question, the remainder of the sentence being backgrounded. In many cases, part of a focused domain, for instance the rightmost constituent, can be deaccented as a consequence of being given, pre-mentioned or somehow salient in discourse or consciousness (see also section 2.4). Selkirk's example of such a case is reproduced in (10). According to her, since the question asks for the whole VP, it is this constituent which is focused. However, since *about bats* is given, it is deaccented. Notice that under "normal" circumstances (when the whole VP is new), *about bats* would get the sentence accent: it is the rightmost constituent, it is embedded deeper than *book* and it is part of an argument. As a result of the deaccenting of *bats*, the accent shifts on *book* (and not on the verb), but still remains on the argument of the verb.

(10)   (*about bats* is "given")
        What did they do?
        Mary [bought a [BOOK]$_F$ about bats]$_F$

An important component of Selkirk's analysis concerns the assignment of F-marking (the subscripted F in (10)) and how it projects. The focus-marking

originating from the pitch accent on *book* can project on the whole verbal phrase, making the sentence a felicitous answer to the question. Notice that it does not say anything about the possible presence of additional accents in the VP domain. This is a crucial point to understanding our data, discussed in the next section. However, an accent on an adjunct or a modifier cannot project in this way, at least according to Selkirk's rules, reproduced here.

(11)   Basic Focus Rule (Selkirk 1995:555)
        An accented word is F-marked

(12)   Focus Projection
        (a) F-marking of the head of a phrase licenses the F-marking of the
            phrase

        (b) F-marking of an internal argument of a head of a phrase licenses the
            F-marking of the head.

In sum, Selkirk's projection rules have initiated a productive line of research: the conditions under which a non-normal accent can stand for a larger accent domain.

2.3   Projection and integration

What emerges from the discussion in the preceding sections is that accent domains and their heads are constructed according to certain rules and principles. One instantiation of such principles is Gussenhoven's (1983, 1992) Sentence Accent Assignment Rule (SAAR), formulated in (13) for English and Dutch, but also applicable to German.

(13)   Sentence Accent Assignment Rule (SAAR) (Gussenhoven 1992)
        If focused, every predicate, argument, and modifier must be accented,
        with the exception of a predicate that, discounting unfocused
        constituents, is adjacent to an argument.

This rule says that adjuncts make up their own focus domain, and that an argument-predicate-complex is integrated into one focus domain, in which the accent on the argument counts for the entire domain, at least when they are adjacent, or when only nonfocused material intervenes between them. In (14), stress on *Tangos* projects up to the entire VP (14a), but stress on *Finnland* does not (14b). *Tangos* is a complement of the predicate *komponieren*, but *Finnland* is an adjunct, which, crucially, cannot be interpreted as selected by the verb (see section 4 for discussion). If an argument and a predicate which would normally be integrated in one domain are separated by such an adjunct, they can each form a domain, as well as the adjunct. This is illustrated in (14c), where the domains are signaled with subscripted P (for Phonological Phrase). In our experiment, presented in section 3, we wanted to first test this prediction of SAAR for similar examples.

(14) a.   weil   Halina [P TANGOS komponiert]
          because Halina   tangos   composes

     b.   weil Halina [P in FINNLAND] [P KOMPONIERT]

     c.   weil Halina [P TANGOS][P in FINNLAND] [P KOMPONIERT]

     d.   weil Halina [P MEHRSTIMMIGE] Tangos komponiert

In its function as an indicator of new material, an accent can signal focus on more than just one word. Following Fuchs (1976), Jacobs (1993) and others, we use the term "integration" to denote the construction of accent domains (or Phonological Phrases). Crucial for Gussenhoven (and also for Selkirk) is the observation that stress on an attributive adjective or on an adjunct cannot project in this way. In (14d), stress on *mehrstimmige* 'polyphonic' only denotes a focus

domain on the adjective. The remaining constituents are backgrounded constituents.

The example in (15) shows that unfocused material between an argument and a predicate is invisible, and does not disrupt the formation of an accent domain in the same way as focused material does (compare (14d)). Our experiment also bears on this issue.

(15)   Warum will Malte in Finnland wohnen?
         why   wants Malte in Finland   live
         'Why does Malte want to live in Finland?'

         weil Halina [PhP TANGOS in Finnland komponiert]

In an OT model, accent domain formation can be expressed with the help of universal constraints, like those first proposed by McCarthy & Prince (1993b) and Truckenbrodt (1999) and adapted for German by Samek-Lodovici (2002).

(16) a.   ALIGN: The edge of a syntactic phrase falls together with the edge of an accent domain

     b.   WRAP: A syntactic domain has to be included entirely in an accent domain.

These constraints integrate a predicate and one of its arguments into a single accent domain, and assign separate accent domains on modifiers. Through the mentioning of syntactic constituents, they ensure that syntactic and prosodic constituents fall together by blocking the formation of accent domains not corresponding to syntactic constituents.

Projection (and integration) has been a productive way to investigate the placement of sentence accent, but it is doubtful that letting accents project in a

purely automatic way is the best conceivable approach to sentence stress assignment. In the following, we consider alternatives.

2.4 Givenness

Until now, the discussion has been centered on sentences uttered in a broad focused (or all-new) context, but, as was briefly mentioned in the preceding section, constituents can be salient in the context, or they may have been mentioned before. In the latter case, the question of projection and integration appears in a new light. Selkirk's sentence (10) shows that a constituent which would normally carry the main stress in an all-new sentence can be deaccented because of its given status. In this case, another constituent, argument or head, carries the stress instead, and the question arises as to what are the principles governing this new accent assignment. In Selkirk's account, only arguments and heads can carry such a default accent: adjuncts and non-heads (like adjectives) do not allow projection. In other words, an accent on an adjective or on any other non-head only signals a focus domain not larger than themselves. But what if there is no argument and no head after elimination of the given constituents?

Schwarzschild (1999) shows that, in this case, adjectives do project their accent, if the other constituents of the focus domain are given (see also Büring 2004). Schwarzschild's example, in (17), is comparable to (7), with the difference that the accented constituent is an adjective. Even if Schwarzschild considers only English, his examples are immediately applicable to German.

(17)   (John drove Mary's red convertible. What did he drive before that?)
        He drove [her [BLUE]$_F$ convertible]$_F$. / Er fuhr [ihr [BLAUES]$_F$ Cabrio]$_F$.

Crucial for the placement of the accent is of course the fact that both *Mary* and *convertible* have been mentioned in their respective thematic role in the question. Because of that, they are eliminated as candidates for accenting.

Schwarzschild provides a quasi-OT analysis of the sentence accent. In his model, the constraint AVOID F in (18), which avoids accents, is crucial. The other important constraint influencing accent placement is called GIVENness, and is formulated as in (19). This constraint interprets a constituent without F-marking. The effect of this constraint is that constituents which are not given (new, or whose role in the sentence is not entailed) have to be F-marked.

(18)   AVOID F

        F-mark as little as possible, without violating GIVENness.

(19)   GIVENNESS

        If a constituent is not F-marked, it must be given.

In Gussenhoven's account, by contrast, the given status of a constituent has a different effect on the accent structure of the other constituents. A modifier inserted between an argument and a predicate, if new, forces the forming of three accent domains: one on the argument, one on the adverb and one on the predicate. If the modifier is given, all three constituents are phrased together and only one phrase is constructed.

This section has briefly reviewed different perspectives on the phonology of sentence accent placement regularities. All of them take the syntactic structure as input, as well as the status of the constituents as new or given. These approaches make some crucially different predictions about the prosodic pattern of a sentence consisting of an argument, a modifier and a verb, in this order. In order to check which approach makes the better assumptions, we conducted an experiment, described in the next section. We return to theoretical issues in section 4.

3   Experiment

3.1.1   Hypotheses

We developed an experiment to test some of the predictions of sentence stress placement rules. More specifically, we were interested in the way Gussenhoven's, Selkirk's and Schwarzschild's predictions are implemented in concrete accent patterns. The models presented in section 2 formulate precise hypotheses for the assignment of accents on some constituents as placeholders for larger domains. The hypotheses we wanted to verify were the following (a stands for argument, v for verb, m for modifier, a capital letter stands for a predicted pitch accent and brackets delimit phonological phrases)

Hypothesis 1: a verbal argument is assigned an accent when both the verb and the argument are under focus (new constituents). (Av)

Hypothesis 2: a new modifier is stressed and does not project its accent further, and as a consequence, when the verb is new, it also has to be stressed to signal its status as new. (A)(M)(V)

Hypothesis 3: when the modifier is given, it carries no stress and in this case, the accent of the argument projects up to the verb. (Amv)

We constructed sentences consisting of a verbal argument, an optional modifier and a verb (see an example in (21)). The sentences were inserted in three different contexts: two eliciting VP-focus (20a,b), and one in which the modifier was given (pre-mentioned in the question, (20c)). Context a was followed by a sentence consisting of an argument and a verb, whereas contexts b and c were followed by the same sentences plus a modifier, new or given respectively.

(20)   Robert ist auf dem Wannsee (mit seinem Katamaran) gesegelt.
        'Robert sailed on the Wannsee (with his catamaran).'

a.   All-new (no modifier): Robert ist doch ein richtiger Wassersportler.
     Was hat er diesen Sommer gemacht?
     'Robert is really keen on water sports. What did he do this summer?'

b.   All-new (with modifier): Robert ist ein echter Outdoorfan. Was hat er
     denn bei seinem letzten Trip gemacht?
     'Robert is a real fan of the outdoors. What did he do on his last trip?'

c.   Modifier given: Robert hat doch jetzt einen Katamaran gekauft. Wie
     hat er ihn denn eingeweiht?
     'Robert has now bought a catamaran. How did he christen it?'

Following Gussenhoven (1983, 1992), our examples were constructed in such a way that the modifiers could not be interpreted as selected by the verb, and, as a result, the word order in our sentences is rather marked. In the default word order, such adverbials are located before the argument. An example of the kind of sentences that we wanted to avoid is *seinen Regenschirm im Zug vergessen* 'forget one's umbrella on the train.' In such a construction, the underlined adjunct can be interpreted as selected by the verb, and, as a result, it is preferably located before the verb. In this way, it differs from our examples, since in our sentences, the adjunct would be preferably located before the argument. The marked, or scrambled, word order of our sentences is certainly componential in the explanation of our results, since scrambled constituents, as has been originally observed by von Stechow & Sternefeld (1988), are triggered by a marked information structure, and are thus prosodically marked, as well (see also Féry & Samek-Lodovici 2004 for this issue). We feel, however, that a syntactic explanation in terms of scrambling is not sufficient to explain the puzzle of sentence accent assignment and that a carefully conducted production experiments will ultimately help to elucidate the role played by the prosody.

3.2 Method

3.2.1 Experimental set-up

The context sentences were spoken by a trained speaker in a sound-proof booth and recorded on a Sony TCD-D100 DAT recorder. The speaker was instructed to speak naturally, in a normal tempo. There were 8 sentences, each of them with three contexts, thus 24 utterances altogether. The context sentences were digitized into individual sound files. The sentences consisted of 6 neutral expressions, and 2 idiomatic expressions[3] (see appendix).

A set-up was conceived in which the subjects were in a quiet room with a Power-Point presentation running on a computer. The experimenter gave brief initial instructions on beginning and ending the session. The subjects worked through the presentation in a self-paced manner. It led them through a set of carefully worded instructions, practice utterances, and finally the experiment itself. The instructions made it clear that the aim of the experiment was to elicit natural language. The experiment was under the subjects' control, who had to press the return key in order to start and continue the experiment. The contexts were presented both auditorily and visually. The subject heard and read first a context, and had to read aloud the sentence appearing on the screen, which was an adequate answer to the question they had just heard. After hitting the return key, a new stimulus appeared. The whole session was recorded on a DAT recorder.

The stimuli were presented in a randomized order and were interspaced with fillers from another experiment bearing on elicitation of natural language, but with a different pattern.

---

[3] Idiomatic expressions are such that their meaning has been conventionalized and cannot be reconstructed—or only partly so—from the meaning of their parts.

3.2.2 Subjects

28 subjects took part in the experiment. They were native speakers of Standard German, students at the University of Potsdam, and had no known hearing or speech deficit. All were paid or acquired credit points for their participation in the experiment.

3.2.3 Analysis

Three people (the authors and a student assistant) listened carefully to the 672 sentences (24 sentences x 28 speakers) and established a list of accents. In most cases, all three judges agreed, but sometimes one of the judges had a different opinion from the other two. When this happened, the two authors listened together carefully to the controversial sentences, and could usually decide on the presence or absence of an accent rapidly, but in some other cases (about 50 sentences, or about 7%), we had to examine the pitch track of the sentences, using PRAAT (© Boersma & Weenik). Our criterion for pitch accent was an excursion on the lexical stress of at least 15 Hz for females and 20 Hz for males.

3.3 Results

1) First, the sentences consisting of an argument and a predicate (sentences a) have a single accent on the argument in the overwhelming majority (see Figure 1). Sometimes the verb is also accented (13%), but when the verb is stressed, the argument is stressed, as well. In other words, the argument is always stressed. (Only in a single utterance we had the impression that just the verb was stressed). For the a sentences, our hypothesis 1 (Av) is thus fully confirmed.

2) In the sentences in which the argument and the verb were separated by a modifier, the givenness status of the modifier plays a role. As can be seen from Figure 2, it was stressed in 91% of the sentences when new (b sentences) and only in 58% when given (c sentences).

In both cases, the argument was nearly always stressed (99% and 97%). The interesting and surprising result is the dimension of variation in the accentuation of the verb. Remember that our hypotheses 2) and 3) predict that the verb is stressed when the modifier is stressed (b sentences), and unstressed when the modifier is unstressed (c sentences). These hypotheses could not be confirmed. We found that the conditional probability for an accent on the verb, given an accent on the modifier (in the all-new context), is only 0.163, which is much lower than would be expected if hypothesis 2 were true. The probability of an accented verb, given no accent on the modifier, is relatively low (0.048), as predicted by hypothesis 3. However, a two-sample z test for proportions showed that the difference between these two conditional probabilities is not significant ($z = 1.40$, $p = 0.162$).
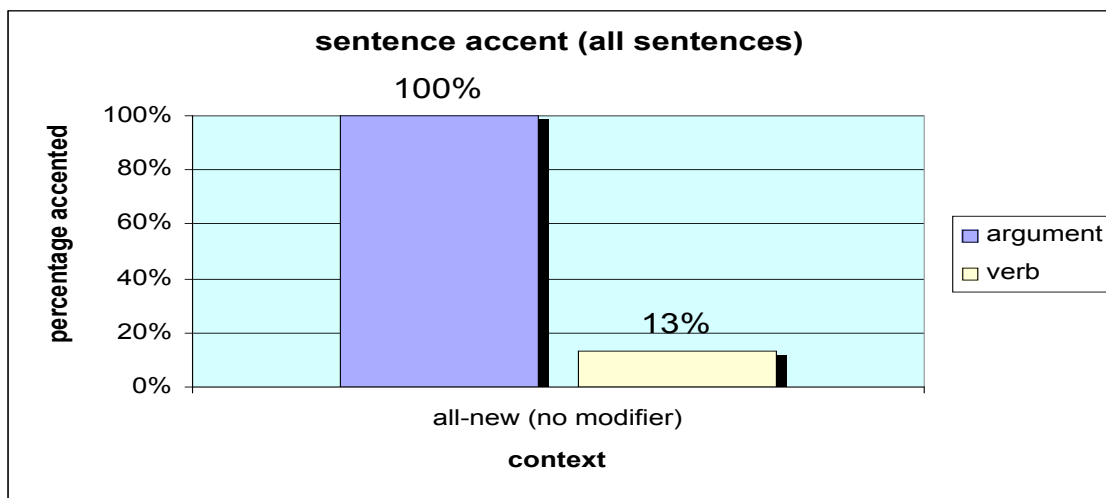
Figure 1
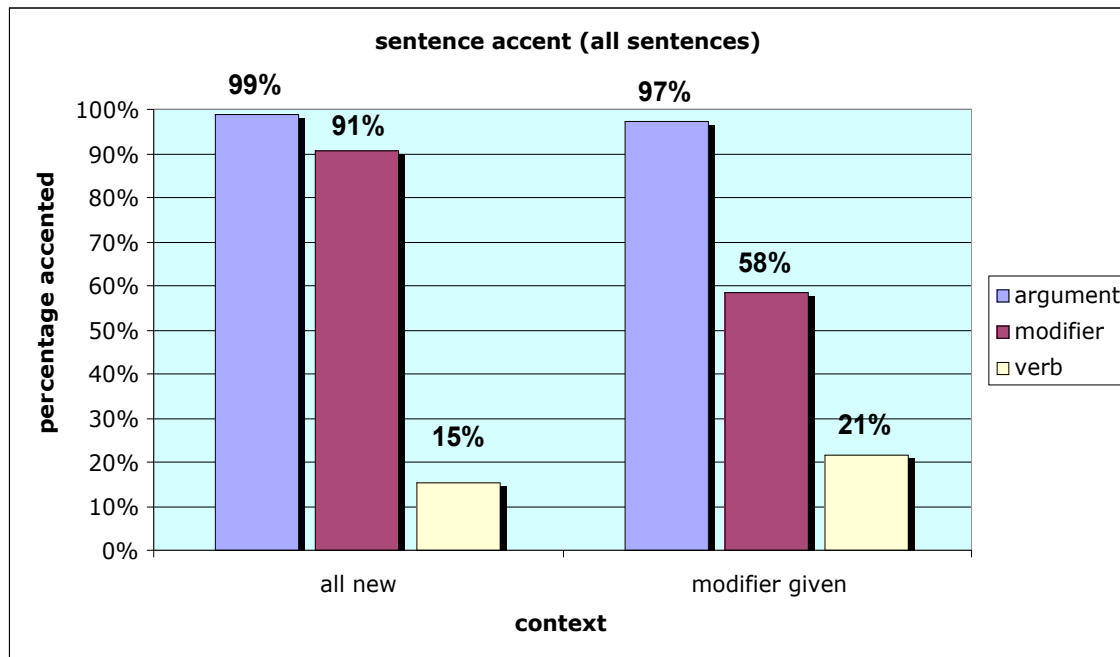
Figure 2

Overall, a new argument is nearly always stressed. When new, the modifier is also mostly stressed, and when given, less so, but still nearly 60% of the time. In sentences with an accented modifier, speakers have thus a tendency to form two accent domains, one on the argument, and one on the modifier plus verb.

The idiomatic and non-idiomatic sentences elicited different results and in the following, we consider them separately.

Non-idiomatic sentences

The non-idiomatic sentences show the effect just described even more clearly. Compare Figures 3 and 4 and Figures 1 and 2 respectively.
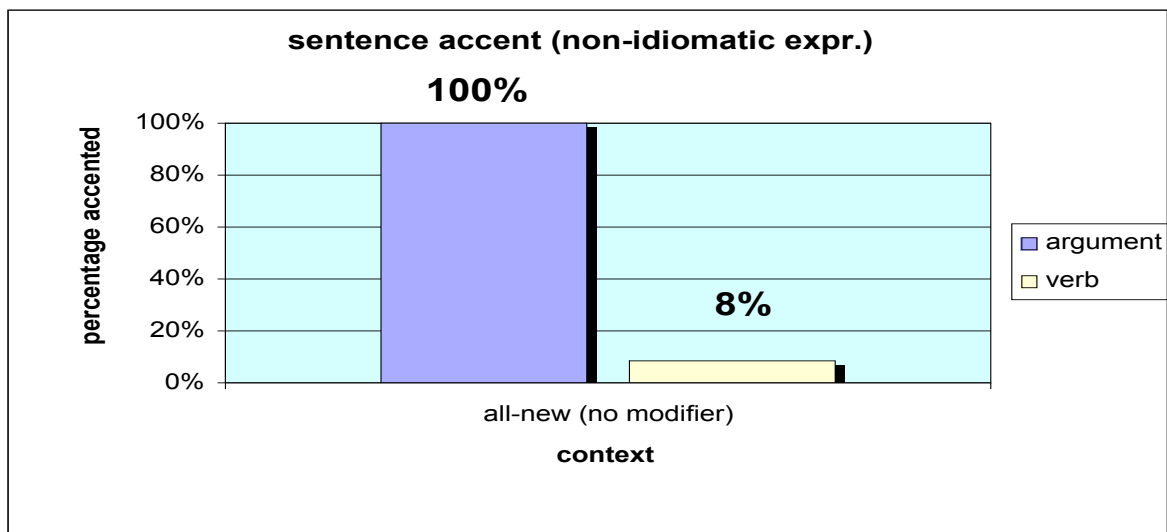
Figure 3
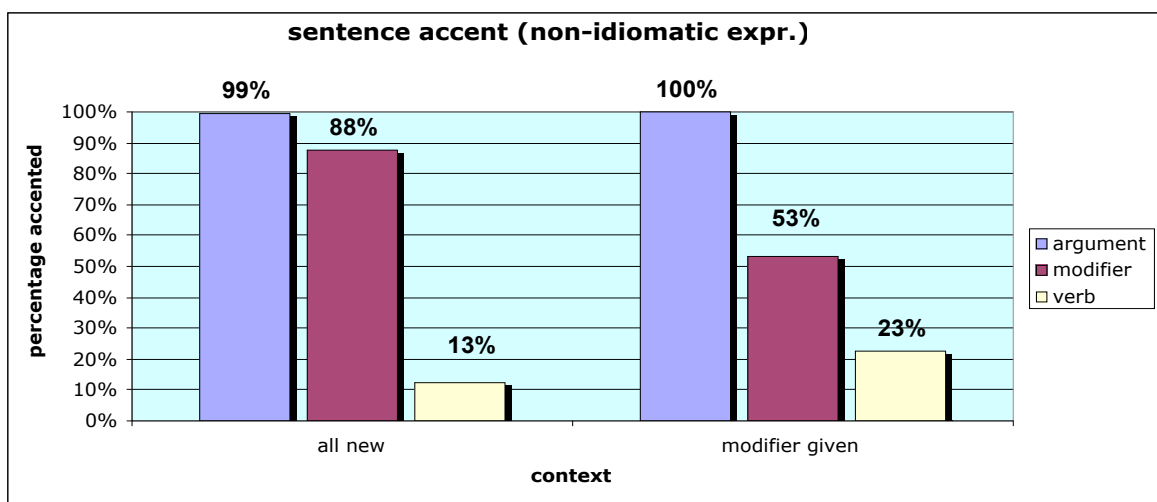
Figure 4

The modifier is stressed in 88% of the cases when new, and 53% when given. The accent of the verb correlates more strongly with this distribution (13% vs. 23%) than in the general overview.[4] And of course, here too, the argument is nearly always stressed (99% and 100%).

---

[4] Most cases of the AMV-pattern (accents on all three constituents) are due to one sentence (sentence 6): 13 out of 20 (for all-new context) and 13 out of 30 (for modifier given) contexts. The verb *schwärmen* 'be mad about' could be non-representative for a reason unknown to us.

Idiomatic sentences

In the idiomatic expressions, the pattern is slightly different. As for sentences with an argument and a verb only, Figures 3 and 5 show that the verb is stressed more than three times more in the idiomatic expressions than in the non-idiomatic ones (29% vs. 8%). This could reflect a property of the sentences chosen for the experiments, or, alternatively, it could also be a general fact about idiomatic vs. non-idiomatic expressions. The last possible explanation—and from an experimental methodology point of view also the most likely one—is to be found in the size of the small corpus used: only two sentences, which means that the result is accidental. More investigations are necessary to settle this issue. A last fact worth mentioning is that the modifier is nearly always stressed, more so when it is new (100%) than when given (75%).

All in all, the idiomatic expressions present a different pattern of stressing, in which the most obvious property is the presence of additional stresses as compared to the non-idiomatic expressions.
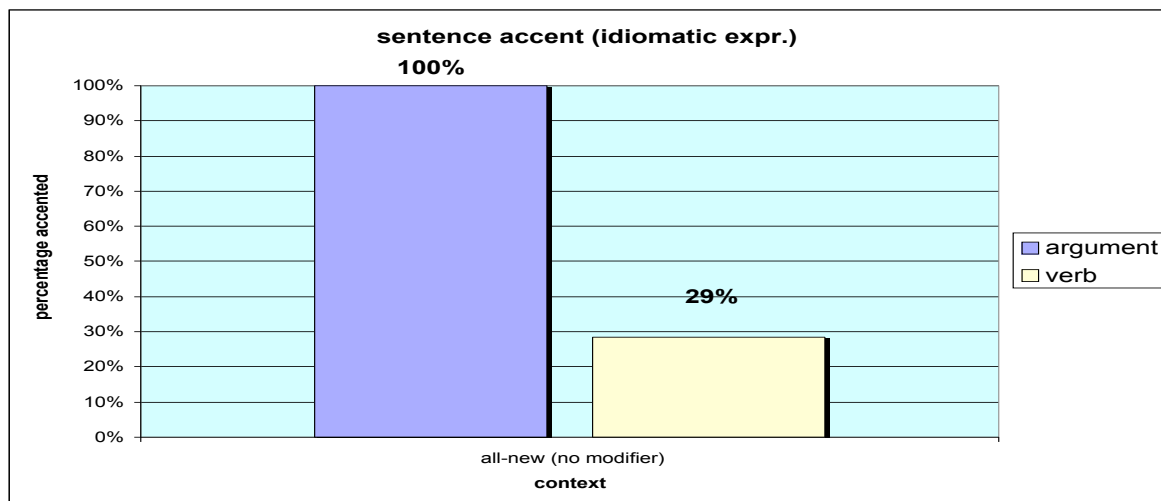
Figure 5
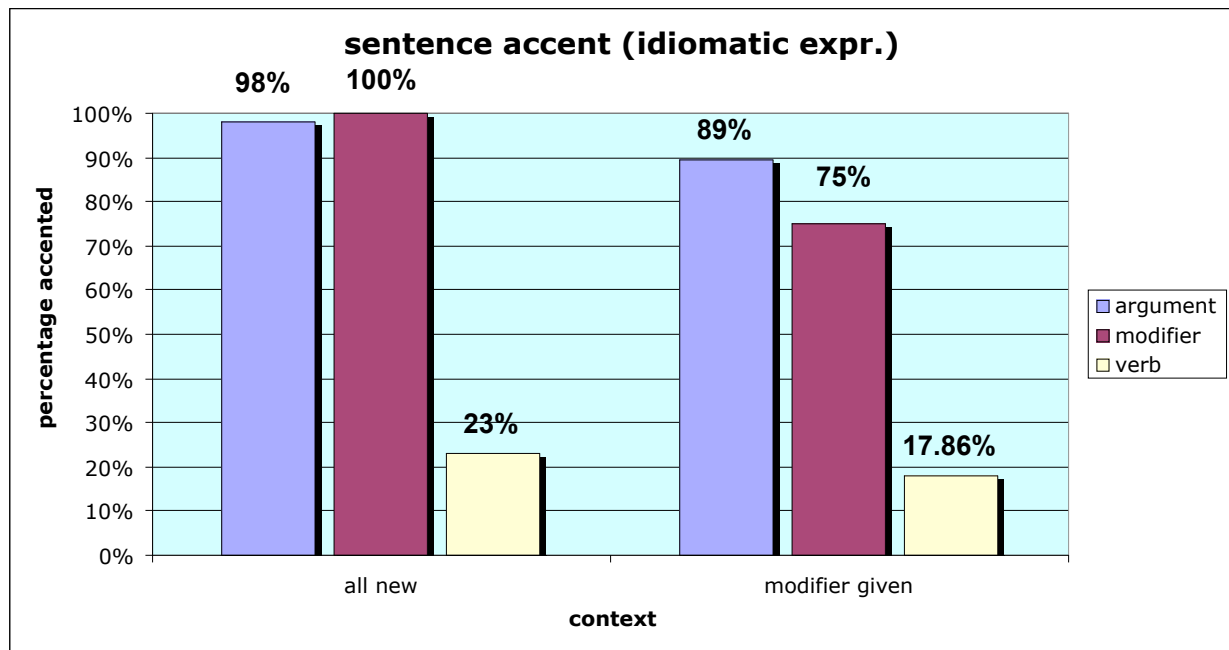
Figure 6

We performed a MANOVA analysis for all sentences, in which the accents on A, M and V were the dependent variables, and the factors were modifier status (new vs. given), and the style of the expressions (idiomatic vs. non-idiomatic).

The MANOVA revealed a significant effect for both modifier status (Wilks' λ = 0.312, p < 0.001) and style (Wilks' λ = 0.003, p < 0.001) in a by-subjects analysis, but only for modifier status in a by-items analysis (Wilks' λ = 0.306, p = 0.006), and not for style (Wilks' λ = 0.699, p = 0.289). However, this may be due to the small sample size, as there were only two idiomatic sentences.

In a subsequent ANOVA, the effect of modifier status on accent was found to be significant for the modifier ($F1_{(1,27)}$ = 45, p = 0.000, $F2_{(1,12)}$ = 6.787, p = 0.023, minF'$_{(1,16)}$ = 5.90, p = 0.028), but not for the argument or verb.

4 Discussion

In terms of the hypotheses formulated in section 3.1, we can sum up our results as follows: (Av) has been experimentally confirmed, and (Amv) only partially. Since in the case of a new modifier, the modifier was tendentially accented, but not the verb, (A)(M)(V) has proven to make wrong predictions and should be replaced by a more adequate representation. A first possibility amounts to allowing a prosodic phrase corresponding to an adjunct to be embedded inside of a larger VP, thus (A(M)v). The accent on the argument could then project to the verb, even though projection takes place across a stressed adjunct. The second possibility projects the stress of an adjunct to an adjacent verb: (A)(Mv). In this configuration, the modifier and the verb are included into one accent domain. The pattern exemplified in (21) is relevant to help us to decide between these two options since it shows that when there is no object, the verb is accented, even though the adjunct is also accented: (M)(V).

(21)   Melina hat [auf der WANDERUNG GESUNGEN]_F (*Melina hat auf der WANDERUNG gesungen) 'During the walk, Melina sang an aria.'

In order to test this pattern, it is again necessary to carefully distinguish between modifiers which seem to be selected by the verb, as in *im Bett liegen* 'to lie in bed', *nach Berlin fahren* 'to go to Berlin', *wegen ihrer Freundin weinen* 'to cry because of one's friend', etc., which have an argumental character, and are readily forming one accent domain with the following verb, and those which are truly sentence modifiers, like *Hans hat auf der Reise geweint* 'Hans cried while on travel' as an answer in a context in which neither *auf der Reise* nor *weinen* are in the background, or *sie hat trotz ihres Unfalls geschlafen* 'she has slept in spite of her accident', in which somebody inquires about her welfare, etc. Only

the latter structures, in which modifier and verb cannot be understood as being included into one meaning unit, allow us to test the modifier-verb pattern.

The stress structure of (21) shows that (A)(Mv) is based on a wrong assumption: when both are in focus, projection from the modifier to the verb is not possible. As was shown in section 2, adjuncts are usually assumed to form their own accent domain, and are not supposed to project their accent to a larger constituent.[5] This is an important insight that we wish to implement into our model, but in an approach without cyclicity, it would have to be abandoned. The second problem related to (A)(Mv) is that it predicts that the prosodic structure can be non-isomorphic to the syntactic structure on a regular basis, an assumption that we do not want to have to defend. Therefore we do not pursue this alternative any further, and instead, explore the first option in some detail.

As mentioned in section 2, Selkirk's bottom-up model can account for the projection of an accented argument to a following verb across an accented modifier. In this respect, it differs from most current accounts of German sentence accent placement, which are based on the linear organization of constituents. In her model, phonetically realized pitch accents project to syntactic domains, and if a certain syntactic domain consists of more than one projecting constituent, its focusing can be obtained by several different accent patterns, depending, among other things, on the givenness of the constituents. In more conventional top-down approaches, it is the focusing of a syntactic domain that is decisive and which leads to a single accent pattern (Cinque 1993, Gussenhoven 1992, Büring 2004, etc.). Envisaged under another perspective, it is the direction of the mapping between accent and focus domain which is responsible for the possible projection of an accent across another accent:

---

[5]   In section 2.4, it was shown that Schwarzschild (1999) and Büring (2004) discuss apparent exceptions. Their examples involve givenness of all alternative bearers of pitch accents and are thus irrelevant for the present discussion.

climbing the tree upwards incidentally permits two or more accent patterns to elicit the same focus domain, whereas descending the tree, only one accent pattern for a single focus structure is ever possible. In sum, Selkirk's model is definitely more indeterminate, as far as sentence accent placement is concerned. This is, on the one hand, a desirable consequence, since it readily accounts for (A(M)v). The accent on the modifier projects no further than to the whole adverbial and the accent on the argument singles out the VP as the focus domain. On the other hand, however, Selkirk's approach also allows an additional accent on the verb, or on another argument, when one is present and, as a result, it is too unconstrained. For this reason, we turn instead to alternatives.

In the remainder of this section, we develop the main lines of an optimality-theoretic account of sentence accent placement. The point of departure of our analysis is that the presence of a modifier between an argument and a verb has no effect on the accent structure of (Av). Recall that Figures 1 and 2 delivered exactly this result: the argument is stressed, but the verb is not, and this regardless of an interfering stressed modifier.

In an OT model, the constraints AVOIDF and GIVENness (Schwarzschild 1999) restrict the number of accents: there is only one accent per domain, and backgrounded constituents may not be accented (22a-b). GIVENness must be ranked very high, but AVOIDF must be in a position in the hierarchy where it can be violated. STRESS-ARGUMENT (see Büring 2001) imposes a stress on the argument. This constraint is higher-ranked than ALIGN-STRESS-R which assigns an accent to the rightmost constituent (22c-d). Finally, the role of WRAP (Truckenbrodt 1999) is to make sure that a syntactic phrase is included as a

whole in a prosodic phrase (22e),[6] and HEAD takes care of the association between a prosodic domain and a head. There must be at least one head in a prosodic domain (22f).[7] In the tableaux shown below, GIVENness does not play any role. ALIGN-STRESS-R has to be lower-ranked than STRESS-ARGUMENT, because, otherwise, candidate b. would win in Tableau 1.

(22) OT constraints

    a.   AVOIDF: Avoid accents.

    b.   GIVENness: Given constituents are not accented.

    c.   STRESS-ARGUMENT: An argument is accented.

    d.   ALIGN-STRESS-R: The rightmost constituent of a phrase is accented.

    e.   WRAP: A syntactic phrase is included in a phonological phrase.

    f.   HEAD: A prosodic constituent has a head.

Tableau 1 shows that the constraints introduced in (22) account straightforwardly for a syntactic domain involving an argument and a verb. Candidate a which violates only AVOIDF and ALIGN-STRESS-R, is optimal because all other competitors violate higher-ranking constraints. Candidate C. also violates AVOIDF, but once more than candidate a.

---

[6] In this paper, we do not discuss the prosodic structure in any detail. Higher prosodic domains may be the result of recursive phonological phrases, or preferably, just higher phrases, like Intonation Phrases.

[7] For our tableaux, we choose WRAP as the constraint responsible for the creation of prosodic phrases, but our data are compatible with an edge-based model, in which the right edge of a maximal projection falls together with the right edge of a maximal projection (McCarthy & Prince 1993b).

Tableau 1. Argument-Verb (*Melina hat* [*eine ARIE gesungen*])

[av]$_F$	HEAD	WRAP	STRESSARG	AVOIDF	ALIGN-STRESS-R
☞ a. (Av)				*	*
b. (aV)			*!	*	
c. (AV)				**!	
d. (av)	*!		*		*
e. (A) (V)		*!	*		*

A structure in which the adjunct is located between the argument and the verb is more problematic. As it is formulated, WRAP prefers a single accent on the whole VP, since, at least on the surface, the VP includes all relevant constituents, thus (Amv), but, as we saw above, (AMv) with two accents is preferred by the speakers. An OT analysis based on ALIGN-XP cannot account for the required AMv pattern, because a high-ranking ALIGN constraint would assign an additional stress to the verb: The maximal projection of the modifier is aligned to its right edge with a prosodic phrase, and, as a result, the verb has to be phrased alone, and has a head. To counter this problem, induced by the presence of an interfering modifier, we propose that the optimality-theoretic model allows a cyclic derivation. The most deeply embedded maximal projections are subject to a first cycle, as illustrated in Tableau 2 for the modifier. Cyclic OT has already been introduced in different guises (see for instance Heck & Müller 2000 for a fully-fledged proposal of successive wh-movement in terms of cyclic OT) though we are not aware of such a proposal accounting for sentence accent.

Tableau 2: Modifier ([*auf der W<small>ANDERUNG</small>*])

[m]_F	HEAD	WRAP	STRESSARG	AVOIDF
☞ a. (M)				*
b. m	*!	*		

At the next cycle, complex projections like the kind of VPs considered in this paper, are computed, but at this level, the adjunct is inserted into the input with its prosodic pattern, as the output of the lower cycle. The pattern (av) is thus a prosodic domain, independent of the presence of an adjunct, as observed in our data. Tableau 3 subsumes both Tableau 1 and Tableau 2 into one, in which M is one domain and (av) another one. Our model allows A to be computed at the lower level, as well, and to enter the higher cycle with its own accent. Merging the verb with its argument could result in a pattern in which the already present accent on the argument is sufficient for the argument-verb complex. As shown in Tableau 3, both the option of computing first the accent of the argument, as well as letting the argument and the verb enter the competition without any previous accent structure would deliver the same accent pattern, and it is thus not possible to decide between the two options. The only difference is that assigning an accent on A on a previous cycle results in a more elaborate prosodic structure. This is shown in Tableau 3 only for candidate a.

Tableau 3: Argument-Modifier-Verb (Melina hat [eine ARIE ([auf der WANDERUNG]) gesungen])

[a(M)v]$_F$ [(A)(M)v]$_F$	HEAD	WRAP	STRESSARG	AVOIDF
☞ a. (A(M)v) ((A)(M)v)				**
b. (a(M)V)			*!	**
c. (A(M)V)				***!
d. (a(M)v)	*!		*	*

5 Conclusion

In this paper, we have shown that one of the assumptions made by standard theories of sentence accent assignment in German which rests on strict linear adjacency between the constituents is not supported experimentally, while others are confirmed. More specifically, we have shown that in sentences consisting of a 'new' argument and a 'new' verb, the sentence accent goes to the argument, as predicted by these theories. The same pattern is also valid when a modifier intervenes between the argument and the verb, regardless of the accentuation of the modifier (the modifier is stressed when new and unstressed when given). This is unexpected under the assumption, mentioned explicitly by Gussenhoven's SAAR, but also present in other models, that a new phrase is created on the verb as soon as the modifier is wrapped in its own, non-projecting phrase.

In order to account for this accent pattern, we have proposed an optimality-theoretic model of German sentence stress assignment that is allowed to apply cyclically. Alternatively, a minimalist model of the prosody-syntax interactions, such as has been proposed by Wagner (2004) for German and by

Ishihara (2003) for Japanese, could turn out to make the best predictions for German. In a first step, the most deeply embedded syntactic projections create accent domains, and in a next step, higher projections are taken into account. These higher projections can then properly contain the accent domains created by the deeper projections. Further investigations will confirm or reject this view of accent domains formation.

Appendix

The question a. elicits an av sentence, where a = Argument and v = Verb, b. elicits amv with m as a new modifier, and c. amv with a given modifier.

1.   Julia ist nach Berlin (mit ihrem Freund) gefahren.
        'Julia went to Berlin (with her boyfriend).'

   a/b. An Wochenenden macht Julia gerne Städtereisen. Was hat sie gestern gemacht?
        'On weekends, Julia enjoys going on city outings. What did she do yesterday?'
   c.   Julia unternimmt ja immer viel mit ihrem Freund. Was hat sie dieses Wochenende gemacht?
        'Julia always does a lot with her boyfriend. What did she do this weekend?'

2   Robert ist auf dem Wannsee (mit seinem Katamaran) gesegelt.
        'Robert sailed on the Wannsee (with his catamaran).'

   a.   Robert ist doch ein richtiger Wassersportler. Was hat er diesen Sommer gemacht?
        'Robert is really keen on water sports. What did he do this summer?'
   b.   Robert ist ein echter Outdoorfan. Was hat er denn bei seinem letzten Trip gemacht?
        'Robert is a real fan of the outdoors. What did he do on his last trip?'
   c.   Robert hat doch jetzt einen Katamaran gekauft. Wie hat er ihn denn eingeweiht?
        'Robert has now bought a catamaran. How did he christen it?'

3   Lisa hat eine Sonne (in ihrem Baumhaus) gemalt.
        'Lisa painted a sun (in her tree house).'

   a/b. Lisa ist so ein kreatives Kind. Was ist ihr denn heute eingefallen?
        'Lisa is such a creative child. What did she do today?'
   c.   In ihrem Baumhaus macht Lisa immer schöne Dinge für uns. Was ist ihr denn heute eingefallen?
        In her tree house, Lisa always makes nice things for us. What did she do today?

4   Melina hat eine Arie (auf der Wanderung) gesungen.
        '(During the walk), Melina sang an aria.'

a/b. Melina ist ein echter Entertainer. Womit hat sie Euch denn diesmal unterhalten?
'Melina is a real entertainer! How did she entertain you this time?'

c.   Ich habe gehört, ihr hattet viel Spaß auf der Wanderung mit Melina. Was hat sie denn gemacht?
'I've heard you had a lot of fun on the walk with Melina. What did she do?'

5   Laura hat bei einer Operation (in der Notaufnahme) zugeschaut.
'Laura watched an operation (in the emergency room).'

a/b. Laura macht doch gerade ihr freiwilliges soziales Jahr. Was hat sie denn gestern so im Krankenhaus erlebt?
'Laura is doing her voluntary year of social service at the moment. What did she see in the hospital yesterday?'

c.   Laura macht bei ihrem Krankenhauspraktikum gerade Station in der Notaufnahme. Was hat sie gestern erlebt?
'At the moment Laura is working in the emergency room as part of her internship at the hospital. What did she see yesterday?'

6   Daniel hat von seiner Heirat (voller Freude) geschwärmt.
'Daniel went on (excitedly) about his wedding.'

a/b. Du hast doch Daniel gestern getroffen. Was hatte er denn zu berichten?
'You met Daniel yesterday. What did he have to tell you?'

c.   Als Du gestern mit Daniel telefoniert hast, klang er so voller Freude. Was hat er denn erzählt?
'When you were talking to Daniel on the phone yesterday, he sounded so happy. What did he tell you?'

7   Jana hat den Braten (beim Meeting) gerochen.
'Jana smelled a rat (at the meeting).'

a/b. Jana war gestern Nachmittag so unfreundlich zu den Kollegen. Was ist denn passiert?
'Yesterday afternoon, Jana was so unfriendly towards her colleagues. What was the matter?'

c.   Jana ist seit dem gestrigen Meeting so schlecht gelaunt. Was ist denn passiert?
Since yesterday's meeting, Jana is so ill-tempered. What was the matter?

8   Der Millionär hat den Löffel (zum falschen Zeitpunkt) abgegeben.
'The millionaire kicked the bucket (at the wrong time).'

a/b. Die junge Blondine hat sich mit dem alten Millionär wohl verrechnet. Letzte Woche ist es passiert:
'The young blonde was mistaken about the old millionaire. It happened last week:'

c.   Manche Dinge passieren einfach zum falschen Zeitpunkt. Letzte Woche ist es passiert:
'Some things just happen at the wrong time. It happened last week:'

References

Berman, Arlene & Michael Szamosi (1972). Observations on sentential stress. *Language* 48, 304-325.

Bierwisch, Manfred (1968). Two critical problems in accent rules. *Journal of Linguistics* 4, 173-178.

Bresnan, Joan (1971). Sentence stress and syntactic transformations. *Language* 47, 257-281.

Bresnan, Joan (1972). Stress and syntax. *Language* 48, 326-342.

Büring, Daniel (2001). What do definites do that indefinites definitively don't?. In: Féry, Caroline & Wolfgang Sternefeld (eds.). Audiatur Vox Sapientae. A Festschrift for Arnim von Stechow. Berlin: Akademie Verlag, 70-100.

Büring, Daniel (2004). Focus Projection and Focus Prominence. Ms. UCLA.

Chomsky, Noam & Morris Halle (1968). The Sound Pattern of English. New York: Harper and Row.

Cinque, Guglielmo (1993). A null theory of phrase and compound stress. *Linguistic Inquiry* 24.2, 239-297.

Fanselow, Gisbert (2004). Cyclic Phonology–Syntax-Interaction: PPT movement in German (and other languages). This volume.

Féry, Caroline & Vieri Samek-Lodovici (2004). Conflicting Foci. Ms. University of Potsdam and UCL, London.

Fuchs, Anna (1976). 'Normaler' und 'kontrastiver' Akzent. *Lingua* 38, 293-312.

Gussenhoven, Carlos (1983). Focus, mode and the nucleus. *Journal of Linguistics* 19, 377-417.

Gussenhoven, Carlos (1992). Sentence accents and argument structure. In: Iggy Roca (ed.) Thematic Structure. Its role in grammar. Berlin: Foris. 79-106.

Heck, Fabian & Gereon Müller (2000). Successive Cyclicity, Long-Distance Superiority, and Local Optimization In: Roger Billerey & Brook D. Lillehaugen (eds.), Proceedings of WCCFL 19, 218-231. Somerville, MA: Cascadilla Press.

Ishihara, Shinichiro (2003). Intonation and interface conditions. Doctoral Dissertation, Massachusetts Institute of Technology.

Ishihara, Shinichiro (2004). Prosody by Phase: Evidence from Focus. Intonation–Wh-scope Correspondence in Japanese. This volume.

Jacobs, Joachim (1993). Integration. In: M. Reis (ed.) Wortstellung und Informationsstruktur. (Linguistische Arbeiten 306). Tübingen. Niemeyer.

Kiparsky, Paul (1966). Über den deutschen Akzent. *Studia Grammatica* VII, 69-98.

Krifka, Manfred (1984). Fokus, Topik, syntaktische Struktur und semantische Interpretation. Ms, Munich.

McCarthy, John & Alan Prince (1993a). Prosodic Morphology I: Constraint Interaction and Satisfaction. University of Massachusetts, Amherst & Rutgers University.

McCarthy, John J. & Alan S. Prince (1993b). Generalized Alignment. In: Geert Booij & Jaap van Marle (eds.) Yearbook of Morphology 1993. Dordrecht: Kluwer, 79-153.

Müller, Gereon (1999). Optimality, Markedness, and Word Order in German. *Linguistics* 37, 777-818.

Newman, Stanley (1946). On the stress system of English. *Word* 2, 171-187.

Prince, Alan & Paul Smolensky (1993). Optimality Theory: Constraint interaction in generative grammar. Univ. of Colorado, Boulder.

Samek-Lodovici, Vieri (2004). Prosody-Syntax Interaction in the Expression of Focus. To appear in *Natural Language and Linguistic Theory*.

Schmerling, Susan (1976). Aspects of English Sentence Stress. Austin, Tex.: University of Texas Press.

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

Selkirk, Elisabeth O. (1995). Sentence Prosody: Intonation, Stress and Phrasing. In: Goldsmith, John (ed.) Handbook of Phonological Theory. Blackwell, 550-569.

Truckenbrodt, Hubert (1999). On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry* 30:2, 219-255.

Stechow, Arnim von & Wolfgang Sternefeld (1988). Bausteine syntaktischen Wissens. Opladen: Westdeutscher Verlag.

Stechow, Arnim von & Susanne Uhmann (1986). Some remarks on focus projection. In: W. Abraham & S. de Meij (eds.) Topic, focus, and configurationality. Amsterdam: John Benjamins.

Wagner, Michael (to appear) Prosody as a Diagonalization of Syntax. Evidence from Complex Predicates. To appear in Moulton, Keir & Matthew Wolf. Proceedings of NELS 34. Stony Brook University.

Caroline Féry and Laura Herbst
Universität Potsdam
Institute of Linguistics
Postfach 601553
14415 Potsdam
Germany
fery@rz.uni-potsdam.de
http://www.ling.uni-potsdam.de/~fery

Prosody by Phase: Evidence from Focus Intonation–*Wh*-scope Correspondence in Japanese[*]

Shinichiro Ishihara

University of Potsdam

Japanese *wh*-questions always exhibit focus intonation (FI). Furthermore, the domain of FI exhibits a correspondence to the *wh*-scope. I propose that this phonology-semantics correspondence is a result of the cyclic computation of FI, which is explained under the notion of *Multiple Spell-Out* in the recent Minimalist framework. The proposed analysis makes two predictions: (1) embedding of an FI into another is possible; (2) (overt) movement of a *wh*-phrase to a phase edge position causes a mismatch between FI and *wh*-scope. Both predictions are tested experimentally, and shown to be borne out.

Keywords: Japanese, wh-*question, prosody, focus intonation,* wh-*scope, cyclicity, phase, Multiple Spell-Out*

1   Introduction

Recently, much attention has been paid to the prosodic properties of *wh*-questions in Japanese and their interaction with syntax and processing (Deguchi and Kitagawa, 2002; Ishihara, 2002; Kitagawa and Tomioka, 2003; Kitagawa and Fodor, 2003; Hirotani, 2003; Ishihara, 2003, among others). It has been claimed that there is a correspondence between the domain of *focus intonation* (henceforth, *FI*)[1] observed in *wh*-questions and the scope of *wh*-questions. It has

---

[1]   Deguchi and Kitagawa (2002) calls it *Emphatic Prosody (EPD)*

also been claimed that this prosody-scope correspondence influences (apparent) syntactic judgments and sentence processing.

In this paper, I will focus on how this prosody-scope correspondence is created. I will claim that prosody is computed cyclically during the course of derivation. Adopting the recent Minimalist framework (Chomsky, 2000, 2001a,b), I propose that cyclic (and hence multiple) application of the so-called *Spell-Out* derives the phonology-semantics correspondence. That is, prosody, the domain of FI in particular, is computed 'phase-by-phase'.

The proposed model makes two predictions. First, it predicts that the cyclic computation of prosody would allow an embedding of an FI into another. Such a pitch contour has not been reported in the literature of Japanese intonation. In fact, standard analyses of Japanese FI (Pierrehumbert and Beckman, 1988; Nagahara, 1994) would not expect such a contour. Second, when a *wh*-phrase is scrambled out of its *wh*-scope, the Multiple Spell-Out analysis predicts that the prosody-scope correspondence will collapse, and result in a mismatch between the FI domain and *wh*-scope. This prediction contradicts the claims made earlier (Ishihara, 2002; Kitagawa and Fodor, 2003), which take the prosody-scope correlation as a principle that Japanese *wh*-questions always comply to. The Multiple Spell-Out analysis proposed here, on the contrary, *derives* the correspondence as a result of the cyclic computation. Under this analysis, the prosody-scope mismatch is a natural consequence of the overt movement of the *wh*-phrase out of its scope. These two predictions are tested experimentally. As we will see, the results of the experiments further support the proposed model.

This paper is organized as follows. In §2, the Focus Intonation–*Wh*-scope Correspondence will be illustrated with actual examples. Then I will propose the Multiple Spell-Out model of FI creation in §3. §4 introduces the two predictions that the proposed model makes. These two predictions are discussed in §5 and §6, respectively, based on the results of the experiments.

2 Focus Intonation–*Wh*-Scope Correspondence (FI=WH)

Japanese *wh*-questions are always accompanied by a focus intonation.[2] Interestingly, the domain of FI exhibits a correspondence to the scope of the *wh*-question, as we will see below. In this section, we will look at some examples showing this phonology-semantics correspondence.

2.1 Focus intonation (FI) in Japanese *wh*-question

Maekawa (1991a,b) showed that Japanese (Tokyo dialect) *wh*-questions exhibit FIs. FIs in Japanese can be characterized by two phonetic phenomena: F_0-boosting on the focalized phrase and the F_0-lowering of the material following the focalized phrase. We will call these phenomena the *P(rosodic)-focalization* and the *post-FOCUS reduction (PFR)*, respectively.

(1) *Focus Intonation (FI) in Japanese*

    a. **P(rosodic)-focalization**

       The F_0 peak of a narrowly focused phrase is raised.

    b. **Post-FOCUS reduction (PFR)**

       The F_0 peaks of the material after the P-focalized phrase is lowered.

A simple illustration of the FI in a *wh*-question is given in (2)[3,4]:

---

[2] There is one more *wh*-construction in Japanese that exhibits FI, namely, the so-called *Mo*-construction (a.k.a. indeterminate construction) (cf. Kuroda, 1965; Nishigauchi, 1990; Shimoyama, 2001; Hiraiwa, 2002). See fn. 20. See also Ishihara (2003) and Kuroda (2004) for discussion on the prosody of *Mo*-construction.

[3] For expository purpose, I will only use lexically accented words in the examples throughout the paper. The location of lexical pitch accent is marked with '´'.

[4] The pitch contours in this examples are recordings of my own voice. All the other pitch contours presented in this paper are obtained from the experiment.

(2) a. *Non-interrogative sentence*

Náoya-ga   **nánika**-o   nomíya-de nónda
Naoya-NOM something-ACC bar-LOC   drank

'Naoya drank something at the bar.'

b. Wh-*question*

Náoya-ga   **náni**-o   nomíya-de nónda **no**?
Naoya-NOM what-ACC bar-LOC   drank Q

'What$_i$ did Naoya drink t_i?'

(2′) a. *Non-interrogative sentence*
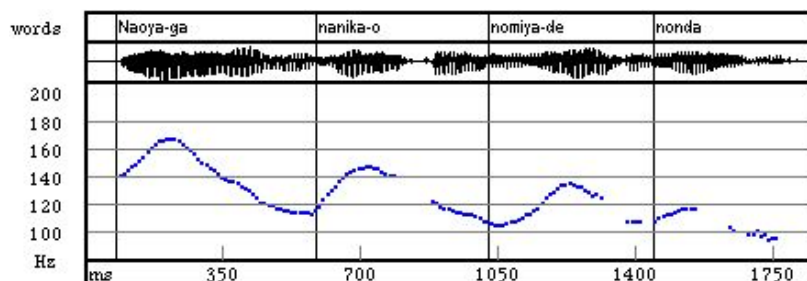
b. Wh-*question*
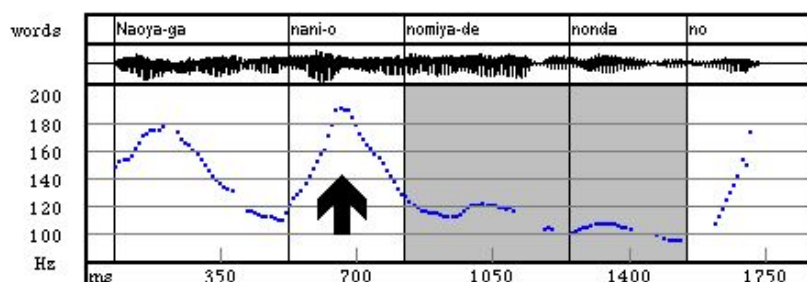
(2a) is a declarative sentence without any narrow/contrastive focus. In this case, the F_0 peaks of the phrases (SUB, OBJ, PP) are all clearly observed.[5,6] On

---

[5] There appears some downstep-like lowering effect on DO and PP in this pitch contour, since they are clearly lower than their preceding phrases and this lowering effect is too large to attribute to time-dependent declination. This lowering effect, however, is not relevant for our discussion, as long as we can observe the contrasts between the declarative sentence and the *wh*-question.

[6] Generally speaking, the F_0-peak of the verb is realized much smaller than XPs (DPs/PPs). I will assume that this is due to *downstep* (a.k.a. *Catathesis*), following Selkirk and Tateishi

the other hand, (2b) is a *wh*-question. The *wh*-phrase DO *nani-o* 'what-ACC' is clearly realized at a higher pitch than the non-*wh*-counterpart in (2a), since the P-focalization on the *wh*-phrase boosts its F_0 peak. In addition to that, the F_0-peaks of the post-*wh*-material, i.e., PP *nomiya-de* 'bar-LOC' and V *nonda* 'drank', are significantly lowered, due to the post-FOCUS reduction.[7]

For the purpose of clarity, I will make one assumption regarding the phonetic nature of P-focalization and PRF, although our main discussion does not hinge on it. Standard analyses of Japanese FI (Pierrehumbert and Beckman, 1988; Nagahara, 1994; Truckenbrodt, 1995, among others) assume that FI is obtained by modifying phonological phrasing, more specifically, by modifying *Major Phrase (MaP)* (a.k.a. *intermediate phrase*) boundaries. A new MaP boundary is created at the focalized phrase while all the MaP boundaries are deleted thereafter. As a result of the restructuring of MaP phrasing, downstep takes place within the newly created large MaP containing the focalized phrase and all the post-FOCUS material. In other words, P-focalization and PFR are captured by the obligatory insertion of a MaP boundary and by downstep, respectively. In this paper, however, I will assume that P-focalization and PFR are pitch-boosting/compression phenomena that are independent of MaP phrasing. This means that I assume that downstep and PFR are different phenomena.[8]

(1991). Downstep is a F_0-lowering phenomenon triggered by H*L pitch accent within a *Major Phrase*. In principle, verbs always receive downstep effect unless they receive a narrow focus. For a detailed discussion on downstep in Japanese, see, among others, Pierrehumbert and Beckman (1988); Selkirk and Tateishi (1991) and Kubozono (1993).

[7] Since the F_0 peaks on verbs are already reduced by downstep (see fn. 6), the effect of PFR may be very small on verbs. Therefore it may often be the case that the expected contrast due to PFR cannot be clearly observed on the verb (e.g., (3′) below). For this reason we will mainly examine the F_0 peaks of non-verbal post-*wh*-phrases.

[8] There are several reasons to take this stance instead of the standard one. Sugahara (2003) shows, for example, that there are cases where MaP boundaries are maintained in the post-focus domain. Even in such cases, however, F_0-lowering is observed, which suggests that PFR is independent of MaP phrasing. See Ishihara (2003) for a more detailed discussion.

2.2   FI–*Wh*-scope Correspondence (FI=WH)

In addition to this prosodic property of *wh*-questions, Deguchi and Kitagawa (2002) and Ishihara (2002) further showed the following property: When a *wh*-question takes matrix scope, its PFR continues until the end of the matrix clause. When a *wh*-question takes embedded scope, its PFR continues until the end of the embedded clause.[9]

Matrix *wh*-question   In the case of a matrix *wh*-question like (3), P-focalization boosts the F_0-peak of the *wh*-phrase, and the PFR compresses the F_0 until the end of the matrix clause, where the question particle *no* appears.

(3)   a.   *Non-interrogative sentence*

Náoya-wa   [Mári-ga   **nánika**-o   nomíya-de nónda to]
Naoya-TOP   Mari-NOM something-ACC bar-LOC   drank that
ímademo omótteru
even.now think

'Naoya still thinks that Mari drank something at the bar.'

b.   Wh-*question*

Náoya-wa   [Mári-ga   **náni**-o   nomíya-de nónda to]
Naoya-TOP   Mari-NOM what-ACC bar-LOC   drank that
ímademo omótteru **no**?
even.now think   Q

'What$_i$ did Naoya still think that Mari drank t_i at the bar?'

---

9   This property is already reported earlier by Tomioka (1997). Thanks to Masa Deguchi for pointing this out to me.

(3′)   a. *Non-interrogative sentence*
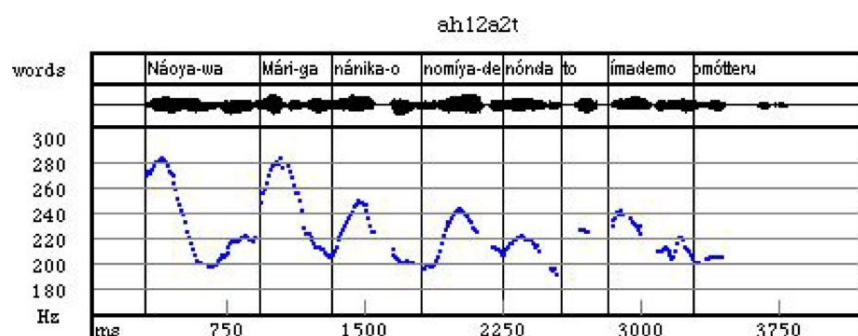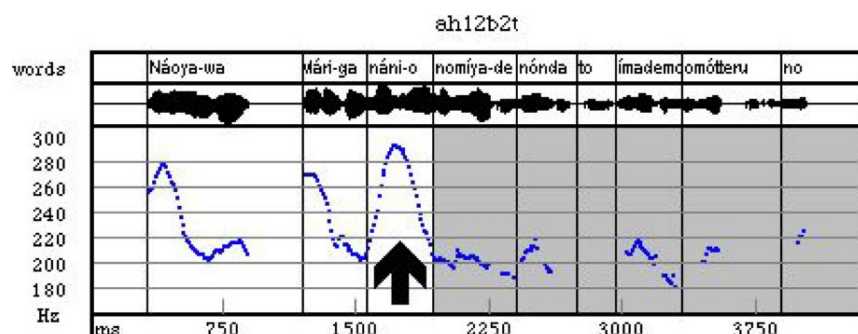
b. Wh-*question*

Indirect *wh*-question   In the case of the indirect *wh*-question in (4), an FI is again observed, but crucially, in a different manner. P-focalization is observed on the *wh*-phrase, as expected. The PFR, however, does not continue until the end of the matrix clause, but stops at the end of the embedded clause, where the embedded Q-particle *ka* appears. In these cases, F_0 exhibits a *pitch reset* phenomenon after the embedded clause: The post-embedded clause material (e.g, *ímademo* in (4b)) shows clear F_0 peaks.

(4)   a. *Indirect Yes/No-question*

Náoya-wa   [Mári-ga   **nánika**-o   nomíya-de nónda ka]
Naoya-TOP   Mari-NOM something-ACC bar-LOC   drank Q
ímademo obóeteru
even.now remember

'Naoya still remembers whether Mari drank something at the bar.'

b. *Indirect* wh-*question*

Náoya-wa [Mári-ga **náni**-o   nomíya-de nónda **ka**]
Naoya-TOP Mari-NOM what-ACC bar-LOC   drank Q
ímademo obóeteru
even.now remember

'Naoya still remembers what$_i$ Mari drank t_i at the bar.'

(4′) a. *Indirect Yes/No-question*

ah11b1t

b. *Indirect* wh-*question*

ah11a1t
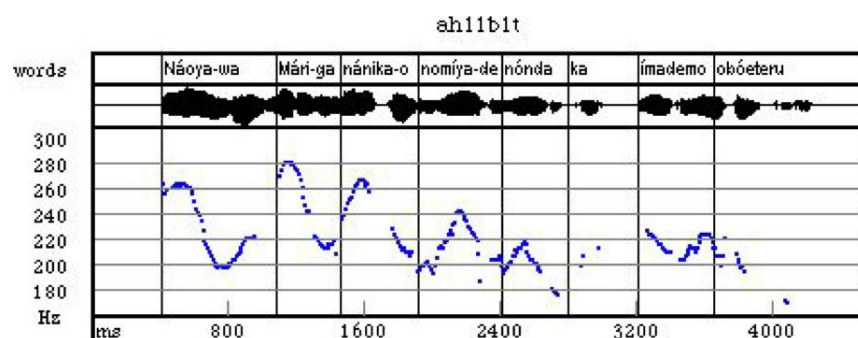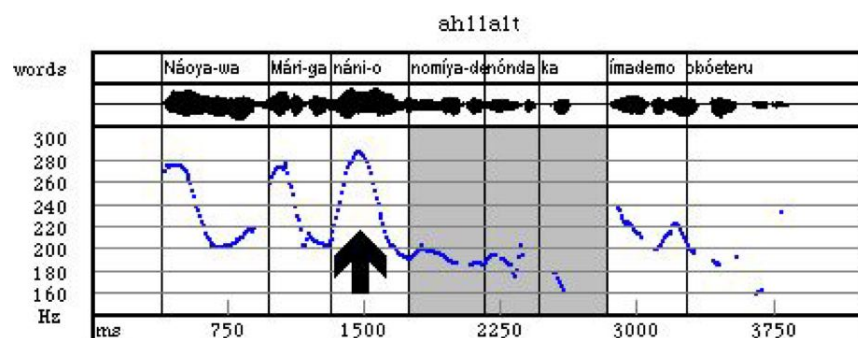
The facts lead us to the following generalization:

(5) *Focus Intonation–*Wh-*scope Correspondence (FI=WH)*[10,11]

The domain of FI corresponds to the scope of a *wh*-question.

---

[10] See Hirotani (2003) for a critical discussion about this generalization.

[11] See also Truckenbrodt (1995, Ch. 4) for a relevant discussion. He claimed that the scope of FOCUS (in the sense of Rooth, 1992) corresponds to the phonological domain at which a focus prominence is assigned.

It should be noted that (5) is just a generalization of the facts we have seen so far. In §6, I will present experimental evidence for the case of *FI–*Wh-*scope Mismatch*, where the FI–*Wh*-scope Correspondence is no longer observed.[12]

The main goal of this paper is to propose a production model that derives this prosody-semantics correspondence, and to present empirical evidence for this model. Although there are many interesting issues regarding the possible effects of prosody on perception or grammatical judgments,[13] I will concentrate on the issues of production in this paper. In the next section (§3), I will present an analysis that accounts for FI=WH, which is based on the recent Minimalist framework (Chomsky, 2000, 2001a,b).

3   A Multiple Spell-Out Account

I propose that FI=WH is a result of the cyclic computation of prosody, which is triggered by the cyclic computation of syntax.[14] This cyclicity in FI creation will be explained in terms of the recent Minimalist framework (Chomsky, 2000, 2001a,b) with the notion of *Multiple Spell-Out*. The syntactic operation *Spell-Out* takes place cyclically at each *phase* in the course of syntactic derivation. My proposal is that prosody, in particular, the domain of FI, is also computed 'phase-by-phase'. In this section, I will present the mechanism of the model I propose.

3.1   Multiple Spell-Out

Multiple Spell-Out is a notion in the recent Minimalist framework proposed by Chomsky (2000, 2001a,b). In this framework, it is proposed that syntactic

---

[12] As we will discuss later (§4.2 and §6), such a case contradicts the empirical claims made earlier by myself (Ishihara, 2002) and by Kitagawa and Fodor (2003).

[13] See Deguchi and Kitagawa (2002); Ishihara (2002); Kitagawa and Fodor (2003); Ishihara (2004) for discussion related to perception issues.

[14] The idea of cyclic phonological computation dates back to Bresnan (1972).

computation is done in a cyclic manner. The unit of this cyclic computation is called the *phase*. At each phase, a certain part of the derivation is transferred (via operation *Transfer*) from the narrow syntax *NS* to two interface levels, Φ and Σ. The phonological part of Transfer, i.e, the operation that transfers the syntactic derivation to the phonological component (NS$\rightarrow\Phi$) is called *Spell-Out*. Since there is more than one phase in a single syntactic derivation, Spell-Out takes place more than once in a cyclic manner during the course of derivation, hence 'Multiple' Spell-Out. The relevant assumptions are listed below.

(6)   *Multiple Spell-Out (Chomsky, 2000, 2001a,b)*

   a.   CPs and *v*P are *phases*.[15]

   b.   When a syntactic derivation reaches a phase (*v*P/CP) in the narrow syntax, the complement of the phase head (i.e., VP/TP) is transferred to the interface levels (Φ/Σ). The phonological part of the Transfer (NS$\rightarrow\Phi$) is called *Spell-Out*.

   [$_{CP}$ (Spec) C [$_{TP}$ (Spec) T [$_{v}$P (Spec) *v* [$_{VP}$ ...]]]]
       $\uparrow$ $\uparrow$ $\uparrow$ $\uparrow$
       phase   Spell-Out   phase   Spell-Out
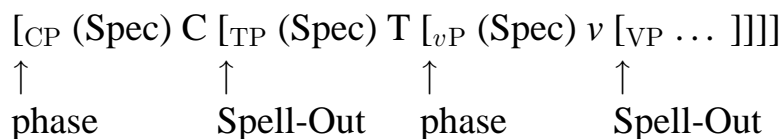
3.2   Proposal

As mentioned at the beginning of the section, the basic claim of the paper is that "FI is created phase-by-phase." In this subsection, I present three relevant assumptions of the cyclic FI prosody model I propose.

FOCUS feature assignment by C   First, we assume that the creation of FI is induced by a FOCUS feature interpreted at the phonological component Φ.[16] I

---

[15] Strictly speaking, only the *v*P of the transitive verb, labeled as *v**P, functions as a phase.

[16] I assume that the FOCUS feature is also interpreted at the semantic component Σ. At Σ, it introduces an alternative set for the focus semantic value (Rooth, 1992). We will not discuss the semantics any further in this paper.

propose that this feature is assigned to *wh*-phrases at the syntactic component by the relevant Complementizers, i.e., Q-particles. Therefore, at that point in a syntactic derivation where a *wh*-phrase is merged to the structure, the *wh*-phrase does not carry a FOCUS feature. It will be assigned to a *wh*-phrase when the relevant Q-particle is merged to the derivation.

(7) *FOCUS feature assignment by C*
     $[_{\text{CP}} [_{\text{TP}} \dots \text{WH}_{\text{FOC}} \dots] \text{C}]$

Timing of FI creation   The FOCUS feature assigned to a *wh*-phrase is interpreted at Φ as soon as it enters into Φ via Spell-Out operation. The FOCUS feature induces P-focalization on the FOCUS phrase and PFR thereafter. Since the Complementizer assigns the FOCUS feature to *wh*-phrases, it is not until C is introduced to the syntactic derivation and a CP phase is formed that the FI creation is induced at Φ.

For example, let us look at the matrix *wh*-question sentence (8), which contains the *wh*-phrase *nani-o* as its object.

(8) $[_{\text{CP}} [_{\text{TP}} \text{Táro-wa} [_{v\text{P}} [_{\text{VP}} \textbf{náni}\text{-o} \quad \text{nónda}] v] \text{T}] \textbf{no}]$
     Taro-TOP   what-ACC drank   Q
     'What did Taro drink?'

When the *v*P phase is created, its Spell-Out domain (VP) contains the *wh*-phrase, but the *wh*-phrase is not yet assigned a FOCUS feature. Thus the FI is not yet created at the Spell-Out of this phase, as in (9a). At the CP phase, the Q-particle *no* is merged to the derivation and assigns a FOCUS feature to the *wh*-phrase. The Spell-Out domain (TP) now contains a FOCUS feature, as in (9b). Hence the FI is created at this Spell-Out cycle.

(9)   a. *vP phase: No FI created*

   [$_{vP}$ [$_{VP}$ **náni**-o nónda] *v*]
            ↑
         No FOCUS feature assigned

   b. *CP phase: FI created*

   [$_{CP}$ [$_{TP}$ Táro-wa [$_{vP}$ [$_{VP}$ **náni**$_{FOC}$-o nónda] *v*] T] **no**]
                                          ↑_____|
                           FOCUS feature assigned by C

FOCUS feature deletion   Lastly, we assume that the FOCUS feature is deleted after the FI is created. This means, once the FOCUS feature is used to create an FI at some Spell-Out cycle, it will not affect prosody created at any later Spell-Out cycle. Let us see how the model works with some examples.

3.3   Examples

The proposed analysis nicely explains the difference in FI realization between the matrix *wh*-question (3b) and the indirect *wh*-question (4b), repeated below. It predicts that the FIs of these two sentences are created at different Spell-Out domains: In the former case, the FI is created at the Spell-Out domain of the matrix CP phase, while in the latter, it is created at the Spell-Out domain of the embedded CP phase. Let us take a closer look at how their FIs are derived.

(3b)   *Matrix* wh-*question: FI created at the matrix CP phase*

   Náoya-wa [Mári-ga **náni**-o   nomíya-de nónda to] ímademo
   Naoya-TOP   Mari-NOM what-ACC bar-LOC   drank that even.now
   omótteru **no**?
   think   Q

   'What$_i$ did Naoya still think that Mari drank t_i at the bar?'
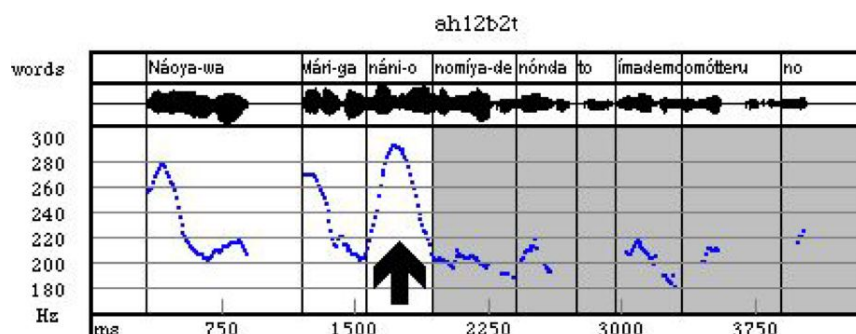
(4b) *Indirect* wh-*question: FI created at the embedded CP phase*

Náoya-wa [Mári-ga **náni**-o nomíya-de nónda **ka**] ímademo
Naoya-TOP Mari-NOM what-ACC bar-LOC   drank Q   even.now
obóeteru
remember

'Naoya still remembers what$_i$ Mari drank t_i at the bar.'
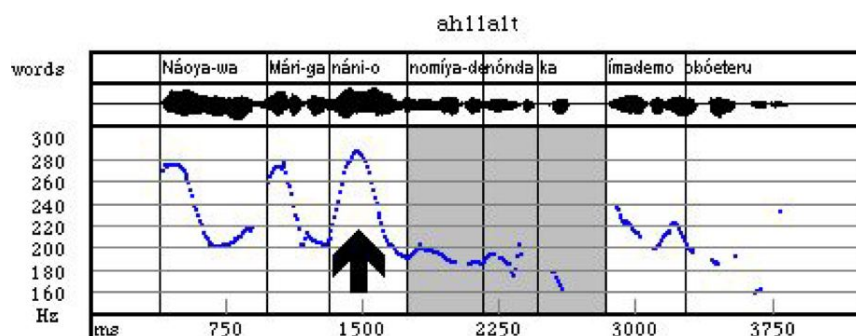
Matrix *wh*-question   In the case of matrix *wh*-questions, the *wh*-phrase is P-focalized, and the PFR after the *wh*-phrase continues until the end of the sentence. (In the examples hereafter, P-focalization is indicated by ⃞box⃞, and PFR by underline.)

(10) *Matrix* wh-*question*
      $[_{CP} \ldots \alpha \ldots [_{CP} \ldots \beta \ldots \boxed{\text{WH}} \ldots \gamma \ldots] \ldots \delta \ldots Q]$

At the embedded CP phase (11a)[17], the *wh*-phrase is not yet assigned a FO-CUS feature, since the Q-particle is not yet merged to the derivation. Therefore its Spell-Out domain, the embedded TP, does not contain any FOCUS feature. Since there is no FOCUS feature, no FI is created at Φ at this point of the derivation, as in (11b).

(11)   a. *Embedded CP phase*

$[_{CP} [_{TP} \dots \beta \dots WH \dots \gamma \dots]]$ (No FOCUS assignment)

   b. *Output at Φ*

$[_{TP} \dots \beta \dots WH \dots \gamma \dots]$ (No FI)

The derivation continues to the matrix CP phase. A Q-particle is merged as the matrix C, and assigns a FOCUS feature to the *wh*-phrase, as in (12a). As a result, the Spell-Out domain, the matrix TP, now contains a FOCUS feature. Accordingly, an FI is created at Φ: The *wh*-phrase is P-focalized, and the PFR applies to the all the post-*wh*-phrases, as in (12b). Since the FI is created at the matrix CP phase, its Spell-Out domain, i.e., the matrix TP, serves as the domain of FI. This means that the PFR domain contains the post-*wh*-phrase material in the embedded CP (γ) as well as the one in the matrix CP (δ).

(12)   a. *Matrix CP phase*

$[_{CP} [_{TP} \dots \alpha \dots [_{CP} \dots \beta \dots WH_{FOC} \dots \gamma \dots] \dots \delta \dots] Q]$

FOCUS assignment

   b. *Output at Φ*

$[_{TP} \dots \alpha \dots [_{CP} \dots \beta \dots \boxed{WH} \dots \gamma \dots] \dots \delta \dots]$

FI creation

At this point, there are a few more elements that have not been transferred to Φ, namely, phrases in the Spec,CP (if any), and the phase head, i.e., Q-particle.

---

[17] Although I will omit *v*P phases for brevity, the explanation presented here for the CP phase not headed by a Q-particle holds for *v*P phases as well. See also fn. 20 for discussion of the *Mo*-construction, in which *v*P phase seems relevant.

I will assume that there is another Spell-Out operation that applies to the root of the derivation, which I will call the *root Spell-Out*. The Spec,CP and the phase head C are transferred to Φ at the root Spell-Out.

Since Q-particles are phase heads and appear outside the Spell-Out domain (i.e., TP), the proposed analysis would predict that they are not to be inside the domain of PFR. In reality, however, these particles seem to be within the PFR domain.[18] I suggest that this is because they do not behave as Prosodic Words by themselves and have no ability to create a new prosodic boundary at any level (Minor Phrase, Major Phrase, or Intonation Phrase). Hence, they are always integrated into the prosodic phrase of the preceding phrase (i.e., verbal complex). Their F_0 is therefore always dependent on that of the preceding phrase.

Indirect *wh*-question   In the case of indirect *wh*-questions like (4b), FI is only observed within the embedded CP. After the embedded CP, a pitch reset is observed. The matrix material after the embedded CP (δ) is outside the FI domain.

(13) *Indirect* wh-*question*

$$[_{\text{CP}} \ldots \alpha \ldots [_{\text{CP}} \ldots \beta \ldots \boxed{\text{WH}} \ldots \gamma \ldots \text{Q}] \ldots \delta \ldots]$$

$$\uparrow$$

Pitch reset

At the embedded CP phase (14a), the Q-particle assigns FOCUS to the *wh*-phrase. When Spell-Out applies to the derivation, the sister of the Q-particle, i.e., TP, is transferred to Φ. Since this Spell-Out domain contains a FOCUS

---

[18]   In the case of the matrix Q-particles like the one in (12a), a question-final rising intonation is normally observed on the Q-particle. Therefore it looks as if they were outside the PFR domain. This rising intonation, however, is not a property of the Q-particle itself, but rather a utterance-final boundary tone that is realized on the final mora of the utterance. If a non-monomoraeic Q-particle *ndai* (cf. Yoshida, 1998) is used, for example, the rising intonation is realized on the last mora of this particle, instead of the beginning of this particle. Even if the Q-particle is omitted (cf. Yoshida and Yoshida, 1996), the rising intonation is still observed on the last mora of the verbal complex. See also fn. 19 about the Q-particle in the embedded clause.

feature, an FI is created: P-focalization on the *wh*-phrase followed by the PFR of the post-FOCUS material (γ), as in (14b).[19]

(14) a. *Embedded CP phase*

$$[_{CP} \ [_{TP} \ldots \ \beta \ldots \ WH_{FOC} \ldots \ \gamma \ldots \ ] \ Q \ ]$$

FOCUS assignment

b. *Output at* Φ

$$[_{TP} \ldots \ \beta \ldots \ \boxed{WH} \ldots \ \gamma \ldots \ ]$$

FI creation

Note that the FOCUS feature is deleted after the FI is created. At the matrix CP phase, therefore, no more FI is created, as in (15). Since the FI is created at the earlier Spell-Out cycle, it does not affect the material introduced at the matrix cycle (α, δ). Accordingly, a pitch reset is observed after the embedded CP.

(15) a. *Matrix CP phase*

$$[_{CP} \ [_{TP} \ldots \ \alpha \ldots \ [_{CP} \ldots \ \beta \ldots \ \boxed{WH} \ldots \ \gamma \ldots \ Q \ ] \ldots \ \delta \ldots \ ] \ ]$$

(No more FOCUS assignment)

b. *Output at* Φ

$$[_{TP} \ldots \ \alpha \ldots \ [_{CP} \ldots \ \beta \ldots \ \boxed{WH} \ldots \ \gamma \ldots \ Q \ ] \ldots \ \delta \ldots \ ]$$

$\uparrow$

Pitch reset

In sum, the FI for a *wh*-phrase is created at the phase whose head is the Q-particle that binds the *wh*-phrase. When the Q-particle is the matrix C (i.e.,

---

[19] In this case again, the Q-particle, which is outside of the Spell-Out domain of the embedded CP phase, appears to be contained in the PFR domain. In my experimental data, there were cases where a sharp F_0 rise is observed on Q-particles, which could potentially be analyzed as a beginning of a new phonological phrase. My impression was, however, that the occurrence of this rise were inconsistent enough to conclude that Q-particles always start a new phonological phrase. Therefore I will assume here that this rise is some sort of boundary tone at the end of the PFR domain. I will leave the investigation of this rise for future research.

when the sentence is a matrix *wh*-question), the FI is created at Spell-Out of the matrix CP phase (i.e., the matrix TP). When the Q-particle is the embedded C (i.e., when the sentence is an embedded *wh*-question), the FI is created at the Spell-Out of the embedded CP (i.e., the embedded TP). Accordingly, the domain of FI corresponds to the scope of the *wh*-question. FI=WH is a result of the cyclic computation of FI.

It should be noted that there is no direct interaction between the phonological and the semantic component during this process. FI=WH is not a result of the direct interaction between phonology and semantics. It is rather the result of the cyclic syntactic computation. One advantage of this model is that the phonological process is as simple as possible. The phonological component only looks for a FOCUS feature each time a new syntactic material is transferred via Spell-Out. When it finds one, it immediately creates an FI. The phonological component is completely indifferent to the semantic scope. Note that the phonetic rules to create an FI are also simple: boosting the F_0 peak of the phrase bearing a FOCUS feature, and lowering everything thereafter. It does not involve specifying where PFR ends. The end point of PFR is automatically derived, since PFR only applies to a relevant Spell-Out domain, not to the whole sentence.[20]

---

[20] One might wonder if there is a case in which an FI is created at a *v*P phase. In the so-called *Mo*-construction (Shimoyama, 2001) (the *indeterminate construction* of Kuroda, 1965), FIs can be found between *wh*-phrase and the particle *mo*, which may appear after C, Verb, or Case-markers. (i) is an example where *mo* attaches to *v*P.

(i) Mo-*construction*
    Mári-wa [*v*P **náni**-o   nomíya-de nómi]-**mo** si-**nákat**-ta
    Mari-TOP   what-ACC   bar-LOC   drink -MO   do-NEG-PST
    'For no x, Mari drink x at the bar.'

This suggests that *v*P and DP are also phases and an FI can be created at their Spell-Out.

4   Two Predictions

The Multiple Spell-Out account proposed in the previous section derives FI=WH as a consequence of cyclic computation in syntax, namely, Multiple Spell-Out. Because of this cyclic property, the proposed analysis makes two interesting predictions. These two predictions are not expected in other possible analyses for FI and for FI=WH phenomena in Japanese. These two predictions are experimentally tested. As we will see, the results strongly support the Multiple Spell-Out analysis.

4.1   Prediction 1: FI embedding

FIs are created cyclically under the Multiple Spell-Out model, it would be possible for a single derivation to create two FIs at different Spell-Out domains. We can therefore make the following prediction:

(16)   *FI embedding*
       When there are two independent WH-Q dependencies with different scopes, an FI is embedded into another.

       [WH1 ... [... WH2 ... α ... Q$_{emb}$] ... β ... Q$_{mat}$]

The resulted contour would realize an FI at the matrix CP (between WH1 and Q$_{mat}$) which contains 'residues' of another FI that are created at the embedded CP (between WH2 and Q$_{emb}$). WH2 would be first P-focalized at the embedded CP phase, and then reduced by PFR at the matrix cycle induced by WH1. Also the post-WH material α would exhibit the PFR effects of both FIs, while the post-embedded CP material β would only show the PFR effect of the matrix FI.

Such a pitch contour has never been reported for Japanese, at least to my knowledge. If such a contour is in fact observed, standard analyses of Japanese

FI (Nagahara, 1994; Truckenbrodt, 1995; Selkirk, 2000; Sugahara, 2003, among others) would require some modifications. As mentioned earlier (§2.1), they assume that FI is obtained by restructuring MaP phrasing. The FI embedding would then be analyzed as an embedding of a MaP into another. Such a prosodic phrasing structure would violate the Non-recursivity of the Strict Layer Hypothesis (Selkirk, 1984; Nespor and Vogel, 1986).[21] Also, Selkirk's (2003) claim that a (contrastive) focus is always associated with prominence at the Intonation Phrase (IP) level would not hold in the FI embedding case, because the realization of the matrix focus (WH1 in (16)) and that of the embedded focus (WH2) are expected to be different: The embedded focus would have a more compressed realization than the matrix focus. In §5, I present and discuss the result of the experiment conducted to test this prediction. In the next subsection, we consider the second prediction.

4.2   Prediction 2: FI–WH Mismatch (FI≠WH) due to movement

The second prediction of the Multiple Spell-Out analysis is related to syntactic movement. So far, we have only seen examples where the *wh*-phrases stay in-situ. In all these examples, we observed FI=WH. Once the *wh*-phrase overtly moves outside the Spell-Out domain via so-called 'edge' position of phases (i.e., the specifier of the phase head), however, the Multiple Spell-Out model expects a different FI than what we have seen so far.

    If a *wh*-phrase moves out of the *wh*-scope phase, by moving to the 'edge' positions in a successive cyclic manner, it will be excluded from the Spell-Out domain of each phase. As a result, the creation of an FI will be postponed to

---

[21] There have been, however, cases reported in the literature that violate Non-recursivity (Selkirk, 1993; Truckenbrodt, 1995). Therefore if FI embedding is in fact the case, it could serve as evidence for MaP embedding. See Kubozono (2004) for the recursive structure of MaP in Japanese downstep. See also Féry and Truckenbrodt (2003); Truckenbrodt and Féry (2003) for a recursive model of downstep for German.

a later Spell-Out cycle than the Spell-Out of the phase where the *wh*-scope is fixed. As a result, the FI domain becomes larger than the actual *wh*-scope.

(17) *FI–*Wh*-scope Mismatch (FI≠WH)*
      Once the *wh*-phrase bearing a FOCUS feature is moved out of its *wh*-scope via phase 'edge' positions, the FI will be created at the later Spell-Out cycle. As a result, *FI–*Wh*-scope Mismatch (FI≠WH)* will arise.

This prediction is drawn from the following theoretical assumptions:

(18) a. The landing site of Ā-scrambling (including all instances of long-distance scrambling) is Spec,CP (Mahajan, 1994).

      b. Spec,CP is the phase 'edge' position, which is *outside* the Spell-Out domain of this CP phase.

This means that any *wh*-phrase scrambled to a Spec,CP will be excluded from the Spell-Out domain of this CP phase, as in (19).

(19) *Embedded CP phase*

$$[_{\text{CP}} \text{ WH}_{\text{FOC}} [_{\text{TP}} \ldots t_{\text{WH}} \ldots] \text{ C }]$$
$$\qquad \uparrow \qquad\qquad \uparrow$$
$$\qquad \text{phase} \qquad \text{Spell-Out (no FI)}$$
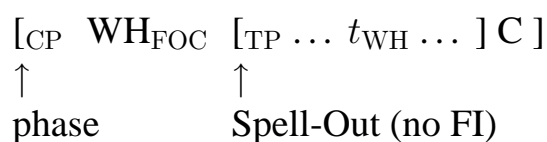
The FOCUS feature of the scrambled *wh*-phrase, then, will be carried to the next phase, i.e., the *v*P phase. As a result, the FI will be created at the Spell-Out of the *v*P phase, namely, VP, which includes not only the embedded clause but also post-embedded-CP phrases (i.e., β in (20)) and the verb.

(20) *Matrix* v*P phase*

$$[_{v\text{P}} \text{ (Spec) } [_{\text{VP}} [_{\text{CP}} \boxed{\text{WH}_{\text{FOC}}} [_{\text{TP}} \ldots t_{\text{WH}} \ldots] \text{ C }] \underline{\beta \ldots \text{ Verb }}] v]$$
$$\quad \uparrow \qquad\quad \uparrow$$
$$\quad \text{phase} \qquad \text{Spell-Out (FI)}$$
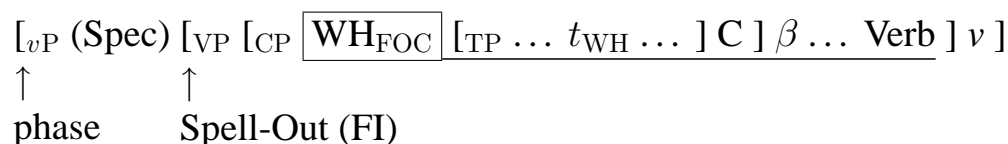
If the *wh*-phrase further moves to a higher position (e.g., Spec,*v*P in (20)) via successive cyclic movement, the FI creation will be delayed further. Semantically, however, the *wh*-phrase is interpreted in-situ, due to the radical reconstruction effect of long-distance scrambling (Saito, 1989). As a result, the domain of FI and the scope of the *wh*-question no longer exhibit a correspondence.

A relevant case can be found in the literature. The example we will examine here is from Saito (1989), in which he showed the radical reconstruction property of long-distance scrambling. In (21), the *wh*-phrase has an embedded scope, regardless of whether the *wh*-phrase is in situ as in (21a), or it is long-distance scrambled to the beginning of the matrix clause as in (21b).

(21)   *Saito's (1989) example: Long-distance-scrambled* wh-*phrase*

    a.   [Mary-ga   [John-ga **dono** hon-o   tosyokan-kara
           Mary-NOM   John-NOM which book-ACC library-from
           karidasita   **ka**] siritagatteiru] koto
           checked.out Q   want.to.know   fact
           'The fact that Mary wants to know [which book]$_i$ John checked out t_i from the library.'

    b.   ? [**dono** hon$_i$-o   [Mary-ga   [John-ga t_i tosyokan-kara
           which book-ACC Mary-NOM John-NOM   library-from
           karidasita   **ka**] siritagatteiru]] koto
           checked.out Q   want.to.know   fact
                                      (Saito, 1989, p. 191–192, ex. (34))

We already saw in (4) that the embedded *wh*-question like (21a) exhibits an FI in the embedded clause, between the in-situ *wh*-phrase and the embedded Q-particle. Now the question is how sentences like (21b) would be pronounced.

If one assumes a non-cyclic model to explain FI=WH, one could generalize FI=WH by stipulating that an FI starts from the *wh*-phrase and ends at the Q-particle that binds the *wh*-phrase. (This was in fact the generalization I made in Ishihara, 2002. See also Kitagawa and Fodor, 2003 for the same claim.) Under such a observation, the expected contour for (21b) would show an FI from

the scrambled *wh*-phrase until the embedded Q-particle *ka*, and a pitch reset thereafter, as illustrated in (22a).

On the other hand, the Multiple Spell-Out model proposed here would predict that the FI is created at the root Spell-Out instead of the embedded CP phase, even though the scope of the *wh*-question is still the embedded CP, due to the radical reconstruction, as in (22b). As a result, we would no longer expect FI=WH. We would rather expect a mismatch between the phonological domain of FI and the *wh*-scope.

(22)   a. *F_0 contour predicted by the generalization in Ishihara (2002)*

$[_{\text{CP}} \boxed{\text{WH}} [_{\text{TP}} \alpha \ldots [_{\text{CP}} [_{\text{TP}} \ldots t_{\text{WH}} \ldots] \textbf{ka}] \beta \ldots]$
$\uparrow$
Pitch reset

        b. *F_0 contour predicted by the multiple Spell-Out model*

$[_{\text{CP}} \boxed{\text{WH}} [_{\text{TP}} \alpha \ldots [_{\text{CP}} [_{\text{TP}} \ldots t_{\text{WH}} \ldots] \textbf{ka}] \beta \ldots]$
$\uparrow$
No pitch reset

If this prediction is borne out, it would pose a challenge to any model assuming direct phonology-semantics interaction to account for FI=WH, because FI=WH no longer holds once the *wh*-phrase is scrambled out of the scope of the Q-particle binding it. If a direct phonology-semantics interaction is assumed to account for FI=*WH*, such a mismatch would not be expected.

In this section, we discussed the two prediction made by the Multiple Spell-Out model. These two predictions were experimentally tested. In the next two sections (§5, §6), the results of the experiments will be presented.

5   Experiment 1: FI Embedding

Let us examine the first prediction, i.e., FI embedding. In this section, I present the result of an experiment, and claim that FI embedding is in fact attested.

5.1   Method

The experiment was conducted using five subjects (four females, AH, CS, CK, NM, and a male, YY), who are all non-linguists brought up in Tokyo or surrounding areas. Stimuli consisting of 32 target sentences (see below for detail) mixed with 104 filler sentences are provided in a pseudo-randomized order (so that two sentences from the same example set are not presented in a row). Each sentence is presented to the subject on a computer screen, one at a time. Subjects are asked first to read the sentence (either aloud or quietly) to understand the meaning of the sentence, and then to read aloud for the recording. Each subject makes 3 recordings of the entire set of stimuli. Each recording uses a different pseudo-randomized order of the stimuli sentences.

5.2   Stimuli

The four sentence types are compared in the experiment. Below is one of the eight stimulus sets used in the experiment:

(23)   *4 sentence types to be examined*

A. **non-WH/WH:** *Indirect* wh-*question*

Náoya-wa [Mári-ga   **náni**-o   nomíya-de nónda **ka**]
Naoya-TOP   Mari-NOM what-ACC bar-LOC   drank Q
ímademo obóeteru
even.now remember

'Naoya still remembers what$_i$ Mari drank t_i at the bar.'

B. **non-WH/non-WH:** *Indirect Yes/No-question*

Náoya-wa [Mári-ga   **nánika**-o   nomíya-de nónda **ka**]
Naoya-TOP   Mari-NOM something-ACC bar-LOC   drank Q
ímademo obóeteru
even.now remember

'Naoya still remembers whether Mari drank something at the bar.'

   C. **WH/WH:** Wh-*question with an indirect* wh-*question*

       dáre-ga   [Mári-ga   **náni**-o   nomíya-de nónda **ka**]
       who-NOM   Mari-NOM what-ACC bar-LOC   drank Q
       ímademo obóeteru   **no**?
       even.now remember Q

       'Who still remembers what$_i$ Mari drank t_i at the bar?'

   D. **WH/non-WH:** Wh-*question with an indirect Yes/No-question*

       dáre-ga   [Mári-ga   **nánika**-o   nomíya-de nónda **ka**]
       who-NOM   Mari-NOM something-ACC bar-LOC   drank Q
       ímademo obóeteru   **no**?
       even.now remember Q

       'Who still remembers whether Mari drank something at the bar?'

(23C) is the FI embedding sentence, which contains one *wh*-phrase in the matrix clause (taking the matrix scope), and another *wh*-phrase in the embedded clause (taking the embedded scope). This sentence is compared with (23D), where the embedded *wh*-phrase is replaced by a non-*wh*-phrase. (23D) would only show an FI at the matrix clause. If FI embedding is possible at all, (23C) would show FI effects at the embedded clause, even though the entire embedded clause is compressed by the PFR of the matrix FI. (23A) and (23B), in which the matrix *wh*-phrase is replaced by a non-*wh*-phrases, are compared with (23C) and (23D), respectively, to make sure that the matrix FI effects are observed in (23C) and (23D).

Among the F_0 peaks in the sentences, those of the following five phrases are measured to examine the FI effects. They are labeled as P1, P2, … P5, respectively.

(24) *Labels of the relevant F_0 peaks*
      [(Non-)WH [… (Non-)WH … α … Verb $C_{[+Q]}$] β … $C_{[\pm Q]}$]
         P1   P2   P3   P4   P5

P1: Matrix *wh*/non-*wh*-phrase   P1 indicates the P-focalization effect at the matrix CP cycle. (If P1 is a *wh*-phrase, it is P-focalized.)

P2: Embedded *wh*/non-*wh*-phrase   P2 indicates the P-focalization effect at the embedded CP cycle. (If P2 is a *wh*-phrase, it is P-focalized.) It also indicates the PFR effect at the matrix CP cycle. (If P1 is a *wh*-phrase, P2 is lowered by PFR.)

P3: Phrase immediately following P2   P3 shows the PFR effects of both the embedded and the matrix CP cycle. (If P1 and/or P2 are *wh*-phrases, P3 is lowered by PFR.)

P4: Embedded clause verb   P4 is not directly relevant to the test.[22] However, since it is the last and the lowest F_0 peak in the embedded clause, it helps us see more clearly the effect of pitch reset expected on P5.

P5: Material immediately following the embedded clause   P5 indicates the PFR effect at the matrix CP cycle, but not the PFR effect at the embedded CP cycle. In other words, P5 indicates the amount of pitch reset after the embedded clause. (If P1 is a *wh*-phrase, P5 is lowered by PFR. If P2 is a *wh*-phrase, P3 and P4 are lowered by PFR, but P5 is not.)

5.3   Predictions

The stimulus set is schematically illustrated in (25). Also, all the expected contrasts are depicted in a graph in (26).

---

[22] Because all the effects expected on this peak are exactly the same as those of P3.

(25)   *Stimulus set (with predicted P-focalization and PFR)*

   A.   *non-WH/WH*

   [P1$_{[-\text{WH}]}$ [… $\boxed{\text{P2}_{[+\text{WH}]}}$ … P3 … P4 C$_{[+\text{Q}]}$] P5 … C$_{[-\text{Q}]}$]

   B.   *non-WH/non-WH*

   [P1$_{[-\text{WH}]}$ [… P2$_{[-\text{WH}]}$ … P3 … P4 C$_{[+\text{Q}]}$] P5 … C$_{[-\text{Q}]}$]

   C.   *WH/WH*

   [$\boxed{\text{P1}_{[+\text{WH}]}}$ [… $\boxed{\text{P2}_{[+\text{WH}]}}$ … P3 … P4 C$_{[+\text{Q}]}$] P5 … C$_{[+\text{Q}]}$]

   D.   *WH/non-WH*

   [$\boxed{\text{P1}_{[+\text{WH}]}}$ [… P2$_{[-\text{WH}]}$ … P3 … P4 C$_{[+\text{Q}]}$] P5 … C$_{[+\text{Q}]}$]

(26)   *Prediction (NB: not an actual result)*[23]
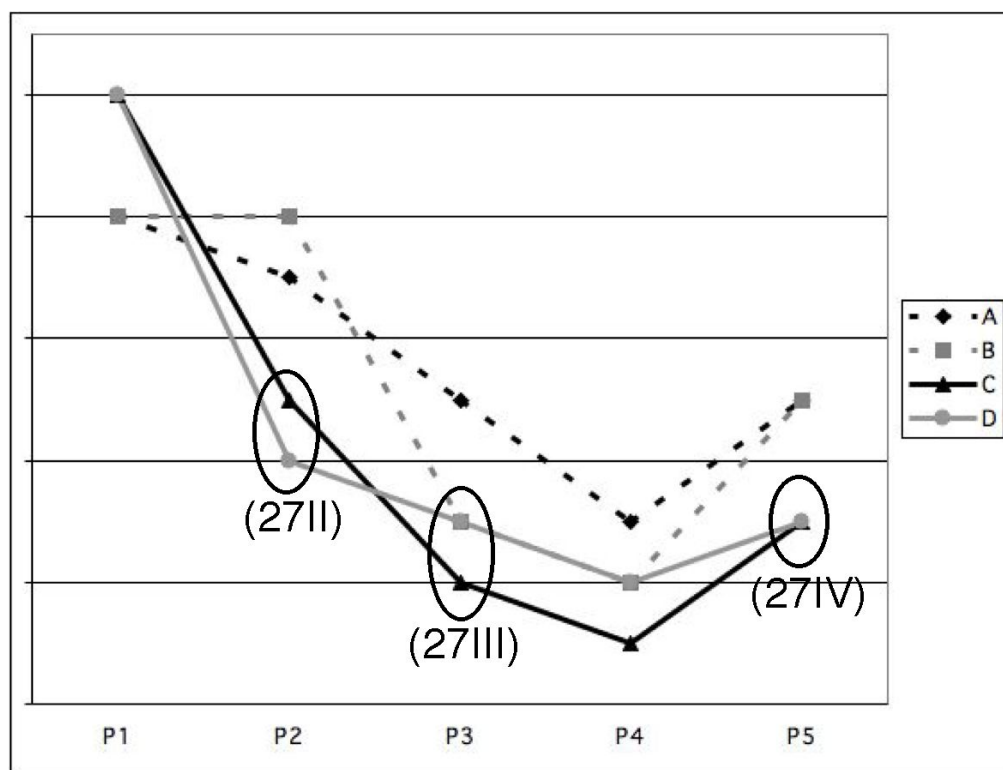
   As mentioned above, the crucial contrasts to be examined is those between
C and D, especially, regarding P2, P3, and P5. First of all, in C and D, all these

---

[23]   This graph simply illustrates the expected contrasts in terms of relative height at each peak
   among the sentence types. No quantitative predictions are made.

peaks are expected to be lowered by the PFR after the matrix *wh*-phrase (P1). It is therefore expected that P2, P3, and P5 are lower in C and D than in B (Prediction (27I)).

Inside this PFR domain, we would expect the difference on P2 and P3 between C and D. P2 is expected to be higher in C due to the P-focalization of the *wh*-phrase (Prediction (27II)); and P3 is expected to be lower in C due to the PFR induced by this *wh*-phrase (Prediction (27III)).

P5, however, is expected to show no difference between the two sentence types, since the embedded FI would not affect this peak (Prediction (27IV)).

(27)   *Crucial predictions*

    I. **P2–5: B $>$ C, D**

      In both C and D, P1 is P-focalized, and P2–P5 are lowered by PFR. Therefore, P2–P5 in C and D are expected to be lower than those in B, where no PFR takes place.

    II. **P2: C $>$ D**

      P2 in C is P-focalized at the embedded CP cycle, while it is not in D. Accordingly, C is expected to be higher than D.

    III. **P3: C $<$ D**

      P3 in C is lowered by PFR at the embedded CP cycle, while it is not in D. Accordingly, C is expected to be lower than to D.

    IV. **P5: C $=$ D**

      P5 in C and D are expected to reach the same height, due to the pitch reset after the embedded CP cycle in D.

5.4   Results and discussion

The results are first analyzed for each subject. Then the data from four of the five subjects (excluding MN's data[24]) are normalized to see if the embedded FI can be observed as a general property among these speakers.[25]

(28)   *Data normalization*

a. Each subject's data is normalized according to the following formula:
$$y = \frac{x - R_2}{R_1 - R_2}$$
where R_1 and R_2 are the reference points calculated independently for each subject.

b. The following two values are chosen as the reference points (R_1, R_2) for the normalization:

- R_1 = Mean value of P1 (F_0-peak on the 1st (non)-*wh*-phrase)

- R_2 = Mean value of P4 (F_0-peak of the embedded verb)

The normalized results are shown in the graph below.

---

[24] In NM's data, not only the expected contrasts, but also other syntax/semantics-related phenomena expected in an utterance (e.g., downstep, utterance final rising intonation for questions) were not attested. The data only showed the time-dependent declination effect. This fact suggests that the subject did not pay sufficient attention to syntax/semantics of the sentences, and read them mere as sequences of words. Such data would not tell us anything important for our purpose.

[25] In this paper, I will only present the normalized data due to space limitations. For the results of the individual subjects and detailed analyses of them, see Ishihara (2003).

(29)   *Normalized Result*
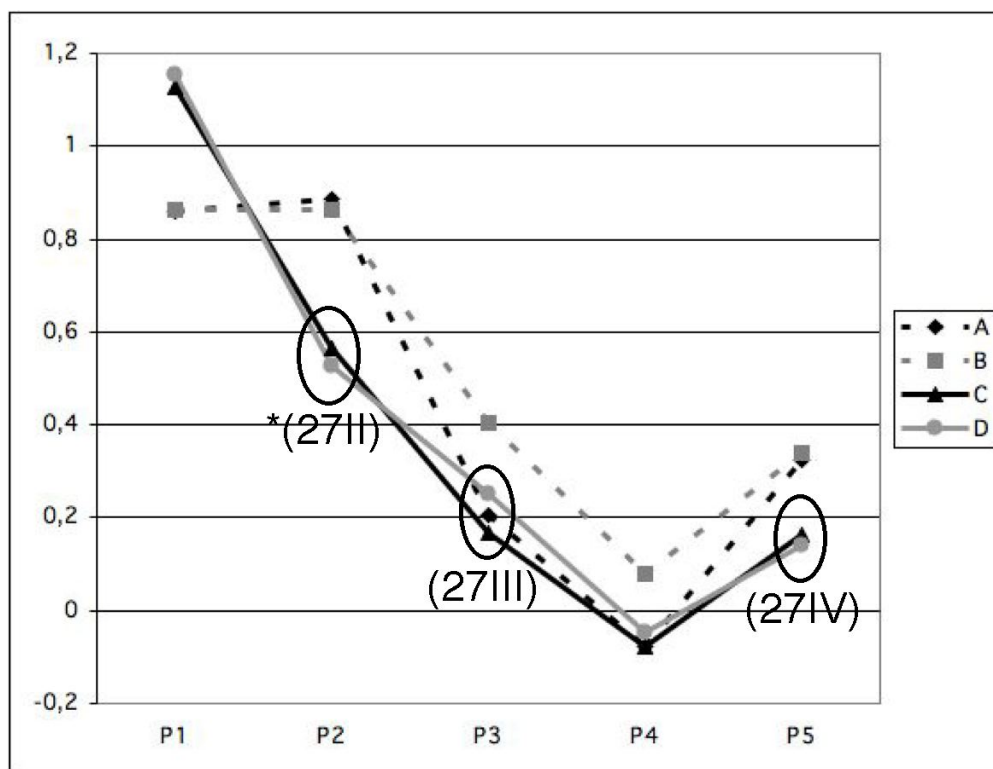
First of all, it is clear from (29) that P2–P5 of C and D are much lower than that of B (i.e., Prediction (27I)). In fact, the contrasts are all statistically significant (p < 0.00001 at all relevant peaks).

Given that P2–P5 are all lowered, we can now compare between C and D to verify the rest of the predictions in (27). The t-test results are shown below:

(30)   *Mean differences between C and D*

Peak	p	Statistically …	Relevant prediction
P2	$= 0.306$	Not significant	*(27II) C > D
P3	< 0.0001	Significant	(27III) C < D
P5	$= 0.231$	Not significant	(27IV) C = D

As shown above, (27III) and (27IV) are supported by the data. There is a statistically significant contrast on P3, showing that P3 is lower in C than in D

(i.e., (27III)). This means that even though P3 is lowered both in C and D by the matrix PFR effect, there is also an embedded PFR effect only in C. This embedded PFR effect is further proved by the fact that P5 in C and D reaches the same point, indicating that there was a pitch reset after the embedded PFR in C. Since the embedded PFR effect in C is limited to the embedded CP, P5, which belongs to the matrix clause, is not affected by this effect. As a result, P5 in C is only affected by the matrix PFR effect, just like in D.

The contrast on P2, however, is not statistically significant. This fact by itself may appear to indicate that there is no embedded P-focalization effect. This lack of expected contrast on P2, however, seems due to the experimental design. As the non-*wh*-counterparts for this position, indefinite pronouns such as *nanika* 'something' and *dareka* 'someone' were used, because they are phonologically minimally different from *wh*-phrases, *nani* 'what', *dare* 'who', etc. I speculate, however, that this similarity made it difficult for the subjects to notice the difference between *wh*-phrase and non-*wh*-counterpart. To my ear, some subjects consistently P-focalized the indefinites as well. As a result, the expected contrast became much smaller than expected. Note that the contrast on P2 is also very small between A and B, as is clear from (29) (p > 0.333). Such a lack of contrast is unexpected, given that the P-focalization effect is clearly attested on P1 (note the difference on P1 between A/B and C/D), where no indefinites were used for the non-*wh*-counterparts. Also note that P2 in B, the F_0 peak of the indefinites, is almost as high as P1. This mean value for P2 is slightly higher than we would expect, given that the time-dependent declination effect would make P2 lower than P1. The speculation about the unexpected P-focalization of indefinites would naturally explain these apparently unexpected facts. Since we do not observe a contrast between A and B, we cannot expect a contrast between C and D either. Given these considerations, the fact that the prediction (27II) is not borne out does not necessarily falsify the analysis.

On the contrary, the other two predictions, (27III) and (27IV), are supported by the result. Given that these contrasts are found within the matrix PFR domain, as (27I) ensures, they clearly indicates the existence of FI embedding.

In this section, we tested the first of the two predictions made by the Multiple Spell-Out model, namely, FI embedding, and discussed the result of the experiment. Although the embedded P-focalization effect was not confirmed by the result, the embedded PFR effect, along with the pitch reset after it, was attested. This result strongly supports the Multiple Spell-Out model proposed in §3. In the next section, we will test the other prediction, namely, FI–WH mismatch.

6   Experiment 2: FI–WH Mismatch (FI$\neq$WH)

In the previous section, we saw that the FI embedding is in fact attested, confirming the first prediction made by the Multiple Spell-Out model. In this section, we will examine the second prediction, namely, the FI–*Wh*-scope Mismatch (FI$\neq$WH). In this experiment, we will examine the pitch contour of Saito's (1989) radical reconstruction sentences like (21b).

6.1   Method

The procedure of the experiment is exactly the same as the one in the FI embedding experiment (see §5.1), except that the number of target sentences is 28 instead of 32, and the number of filler sentences is 108 instead of 104.

6.2   Stimuli

Stimuli are made of 7 sets of four sentence types (28 sentences in total), one of which is given below:

(31) *Stimulus set example*

A. *No scrambling, Non-*wh-*sentence*

Náoya-wa [Mári-ga   **rámu**-o nomíya-de nónda to]
Naoya-TOP   Mari-NOM rum-ACC bar-LOC   drank that
ímademo omótteru
even.now think

'Naoya still thinks that Mari drank rum at the bar.'

B. *No scrambling, Indirect* wh-*question*

Náoya-wa [Mári-ga   **náni**-o   nomíya-de nónda **ka**]
Naoya-TOP   Mari-NOM what-ACC bar-LOC   drank Q
ímademo obóeteru
even.now remember

'Naoya still remembers what$_i$ Mari drank t_i at the bar.'

C. *Scrambling, Non-*wh-*sentence*

rámu$_i$-o Náoya-wa [Mári-ga t_i nomíya-de nónda to]
rum-ACC Naoya-TOP   Mari-NOM   bar-LOC   drank that
ímademo omótteru
even.now think

'Naoya still thinks that Mari drank rum at the bar.'

D. *Scrambling, Indirect* wh-*question*

náni$_i$-o   Náoya-wa [Mári-ga t_i nomíya-de nónda **ka**]
what-ACC Naoya-TOP   Mari-NOM   bar-LOC   drank Q
ímademo obóeteru
even.now remember

'Naoya still remembers what$_i$ Mari drank t_i at the bar.'

(31A) and (31B) are sentences with a canonical word order (i.e., no scram-
bling). (31B) is an embedded *wh*-question, containing a *wh*-phrase and a Q-
particle in the embedded clause.

(31C) and (31D) are the scrambled versions of (31A) and (31B), respectively. (31D) is Saito's (1989) example, where the embedded *wh*-phrase is scrambled to the beginning of the matrix clause.

6.3   Predictions

In this experiment, we are interested in the FI domain of sentences like (31D). What we need to verify is to see whether the FI domain continues after the embedded clause (as the Multiple Spell-Out model predicts) or not (as claimed earlier by Ishihara, 2002; Kitagawa and Fodor, 2003). To test this, we focus on the F_0-peak of the embedded Verb (P1) and that of the phrase after the embedded clause (P2). (In (31): P1 = *nónda*; P2 = *ímademo*)

(32)   *Labels of the relevant F_0 peaks*
$[_{CP}$ ((Non-)WH) ... $[_{CP}$ ... ((Non-)WH) ... Verb Q] α ...]
P1   P2

P1: Embedded clause verb   P1 is inside the embedded CP. Hence it will be lowered if an FI is created either at the embedded CP cycle or at the matrix CP cycle.

P2: Material immediately following the embedded clause   P2   is outside the embedded CP. Hence it will be lowered only if an FI is created at the matrix CP cycle. It will be insensitive to the FI within the embedded CP.

Under the Multiple Spell-Out model, we will have the following predictions for the non-scrambling sentences (A and B) and for the scrambling sentences (C and D), Respectively:

(33)   Non-scrambling sentences: A vs. B

A.   [$_{CP}$ [$_{TP}$ … [$_{CP}$ [$_{TP}$ … Non-WH … P1] C] P2 …]]
                                                          ↑   ↑
                                                    No PFR   No PFR

B.   [$_{CP}$ [$_{TP}$ … [$_{CP}$ [$_{TP}$ …   WH   … P1] **ka**] P2 …]]
                                                          ↑   ↑
                                                        PFR   No PFR

- In B, an FI is created at the embedded CP cycle.

- P1 of B is lower than in A due to the PFR.

- P2 of A and B are of the same height, because P2 should not be affected by the PFR in the embedded CP cycle. Hence, a pitch reset takes place.

(34)   Scrambling sentences: C vs. D

C.   [$_{CP}$ Non-WH$_i$ [$_{TP}$ … [$_{CP}$ [$_{TP}$ … t_i … P1] C] P2 …]]
                                                          ↑   ↑
                                                    No PFR   No PFR

D.   [$_{CP}$ WH$_i$ [$_{TP}$ … [$_{CP}$ [$_{TP}$ … t_i … P1] **ka**] P2 …]]
                                                          ↑   ↑
                                                        PFR   PFR

- In D, an FI is created at the matrix CP cycle.

- P1 of D is lower than that of C.

- P2 of D is also lower than that of C.

6.4   Result and discussion

Data of the four subjects (AH, CS, CK and YY) are normalized.[26]

---

[26] Again, the data of one subject (NM) is excluded from the analysis. See fn. 24.

(35) *Data normalization*

    a. *Formula for normalization:*

$$y = \frac{x - R_2}{R_1 - R_2}$$

   where R_1 and R_2 are the reference points calculated independently for each subject.

    b. *Reference points (R_1, R_2):*

- R_1 = Mean value of P2 (F_0-peak on the phrase immediately following the embedded clause)

- R_2 = Mean value of P1 (F_0-peak on the embedded verb)

The normalized data show the expected results. In the non-scrambled sentences A and B, P1 (the embedded verb) is lowered in B due to the PFR after the *wh*-phrase. The difference between A and B is statistically significant. On P2 (the post-embedded-CP phrase), although there still is a difference between A and B, it is much smaller than the one on P1. It is in fact statistically not significant. This means that in B a pitch reset takes place and the pitch register of the P2 is set back to the non-reduced value. Hence there is no more significant difference on P2.

(36) A vs. B

	A	B	diff.	p
Mean(P1)	0.174	−0.103	0.276	$< .001$
Mean(P2)	1.066	0.971	0.095	$= .257$
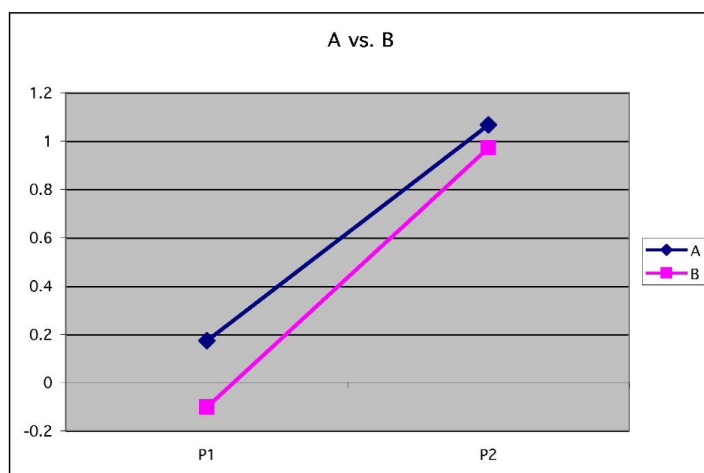
In the scrambled sentences C and D, P1 shows the same result as in the non-scrambled version, as expected. P1 is lower in D than in C due to the PFR. On P2, the differences between C and D are not reduced at all, and in fact, they are still statistically significant. This means that the PFR continues to the matrix material, unlike the non-scrambled version.

(37)   C vs. D

	C	D	diff.	p
Mean(P1)	0.115	−0.185	0.301	< .001
Mean(P2)	1.182	0.780	0.402	< .0001
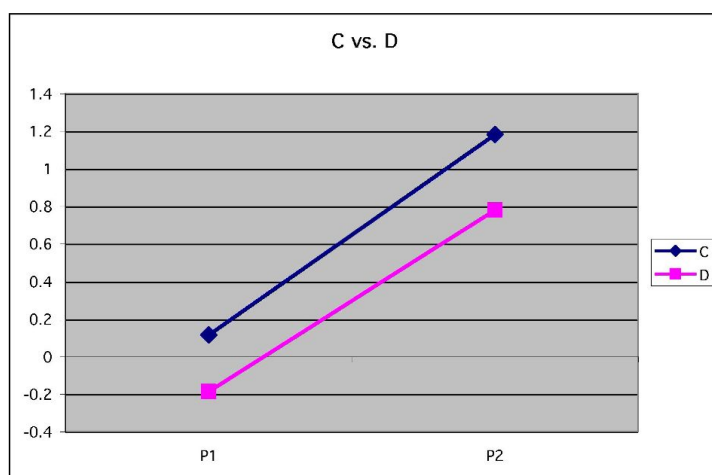
The result of the experiment reinforces the Multiple Spell-Out analysis proposed in this paper, as the prediction is in fact empirically supported. This result also denies the earlier empirical claim in Ishihara (2002) and Kitagawa and Fodor (2003) that the FI is always observed between the *wh*-phrase and the Q-particle.[27]

(38)   a.   * *Observation in Ishihara (2002); Kitagawa and Fodor (2003)*

$$[_{\text{CP}}\ \boxed{\text{WH}}\ \underline{[_{\text{TP}}\ \alpha \ldots\ [_{\text{CP}}\ [_{\text{TP}} \ldots\ t_{\text{WH}} \ldots\ ]\ \textbf{ka}\ ]}\ \beta \ldots\ ]$$

$$\uparrow$$

No PFR

      b.   *Actually attested pitch contour*

$$[_{\text{CP}}\ \boxed{\text{WH}}\ \underline{[_{\text{TP}}\ \alpha \ldots\ [_{\text{CP}}\ [_{\text{TP}} \ldots\ t_{\text{WH}} \ldots\ ]\ \textbf{ka}\ ]\ \beta \ldots\ ]}$$

$$\uparrow$$

PFR

This result is particularly important because it suggests that FI=WH is not a result of the direct phonology-semantics interaction. If it were the case, we would expect an FI only inside the embedded clause in Saito's (1989) example like (31D). FI=WH is rather a result of the cyclic computation, which usually computes the domain of FI and the *wh*-scope at the same phase, unless the syntactic movement creates a mismatch between the phonological domain of FI and the semantic *wh*-scope.

In this section, we examined the second prediction of the Multiple Spell-Out analysis, namely, the FI–*Wh*-scope Mismatch. The result of the experiment

---

[27] An question remains as to why both Kitagawa and Fodor (2003) and I (Ishihara, 2002) acknowledged that (38a) is the correct pitch contour. In fact, I still feel that (38a) is not entirely impossible. It is, however, hard to decide whether this intuition is real and has to be accounted for, because this sentence always involves unnaturalness in judgement (which is in fact the main point of discussion in Ishihara, 2002 and Kitagawa and Fodor, 2003), and maybe also because I may be too sensitive to the FI–*Wh*-scope correspondence to give a naive judgement. If, however, this intuition turns out to be real, there must be some additional mechanism that allows a contour like (38a), because the Multiple Spell-Out model would never allow such a contour. I will leave this question for future research. In this paper, I will take the result of the experiment as the real and correct description of the fact.

presented in this section in fact supports this prediction. When a *wh*-phrase is scrambled out of its *wh*-scope, the FI creation is postponed to a later Spell-Out cycle, and the domain is extended. As a result, the FI domain and the *wh*-scope no longer shows a correspondence. Together with the FI embedding discussed in §5, this experimental result strongly supports the proposed analysis.

7   Conclusion

In this paper, we discussed the Focus Intonation–*Wh*-scope Correspondence (FI=WH) in Japanese *wh*-questions. I proposed that FI=WH is derived by the cyclic syntactic computation and the Spell-Out mechanism.

The Multiple Spell-Out model proposed here is further supported by the results of the two experiments. The first experiment showed that FIs may be embedded when there are two WH-Q dependencies that take different scopes. FI embedding is naturally explained under the proposed model. The second experiment showed that FI=WH breaks down once the *wh*-phrase is scrambled out of its *wh*-scope. The *wh*-scope remains the same if scrambling takes place, thanks to the radical reconstruction effect. The FI prosody, in contrast, is created later in the derivation, namely, at the Spell-Out domain at which the scrambled *wh*-phrase is transferred to the phonological component. As a result, FI$\neq$WH takes place.

This analysis not only explains FI=WH and FI$\neq$WH in Japanese *wh*-questions, but also has further theoretical implications. First, under this analysis, the phonological component computes prosodic information in a cyclic fashion. This means that not only segmental phonological material, but also suprasegmental information such as intonation is computed cyclically phase by phase, and superimposed each time. The FI embedding experiment (§5) suggests that this is in fact the case. If so, it raises further interesting questions such as how the phonological component implements such cyclic suprasegmental in-

formation, how it is realized phonetically, how the cyclic Spell-Out is related to phonological phrasing, etc.

Also, this analysis gives support for the phase 'edge' position. In the current Minimalist framework, phase 'edge' positions are needed at the syntactic component to allow successive cyclic movement. The material (dis)located to this position escapes from Spell-Out at this phase, remaining accessible to the next phase. The FI–WH Mismatch experiment (§6) provides support for this claim. Material moved to this position is in fact spelt-out at a later cycle.[28]

As interesting discussion has already been made recently (see §1 for references), prosody and its impact on syntactic 'judgment' has to be studied more in detail. What is interesting about the prosody of Japanese *wh*-questions is that FIs appear obligatorily in the sentence. The situation is clearly different from non-*wh*-sentences. Since the appearance of focus heavily depends on the discourse and information structure of the sentence, an FI may or may not appear in a non-*wh*-sentence, depending on the context. This does not necessarily mean, however, that *wh*-questions may not have any additional FIs optionally. Some *wh*-question sentences may contain both obligatory and optional FIs. It is therefore important for future research to specify how these 'obligatory' and 'optional' FIs may interact with each other. Such studies would help us understand better how prosody influences syntactic judgments.

Bibliography

Bresnan, Joan W. 1972. On sentence stress and syntactic transformation. In *Contributions to generative phonology*, ed. Michael Brame, 73–107. Austin:

---

[28] The real test case would be where the scrambled *wh*-phrase stops at the specifier of the embedded CP. The Multiple Spell-Out model predicts that the FI will be created at the matrix *v*P phase. This property could potentially be used to distinguish the two types of scrambling. The scrambling to Spec,CP is considered to be A′-scrambling, while the scrambling to Spec,TP is A-scrambling (Miyagawa, 2001). If so, A′-scrambling would create an FI at the matrix *v*P phase, while A-scrambling creates an FI at the embedded CP phase. I leave these further questions for future research.

University of Texas Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: In honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 85–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001a. Derivation by phase. In Kenstowicz (2001), 1–52.

Chomsky, Noam. 2001b. *Beyond explanatory adequacy*, volume 20 of *MIT Occasional Papers in Linguistics*. Cambridge, MA: MITWPL.

Deguchi, Masanori, and Yoshihisa Kitagawa. 2002. Prosody and wh-questions. In *NELS 32: Proceedings of the North East Linguistics Society*, ed. Masako Hirotani, volume 1, 73–92. Amherst, MA: GLSA, University of Massachusetts, Amherst.

Féry, Caroline, and Hubert Truckenbrodt. 2003. Sisterhood and tonal scaling. Unpublished ms.

Hiraiwa, Ken. 2002. Indeterminate agreement and varieties of raising in Japanese. Paper accepted at the 21st West Coast Conference on Formal Linguistics (WCCFL XXI).

Hirotani, Masako. 2003. Prosodic effect on the comprehension of Japanese wh-question. Poster presentation at the Sixteenth Annual CUNY Conference on Human Sentence Processing, May 27–29 2003.

Ishihara, Shinichiro. 2002. Invisible but audible wh-scope marking: Wh-constructions and deaccenting in Japanese. In *Proceedings of the 21st West Coast Conference on Formal Linguistics (WCCFL 21)*, ed. Line Mikkelsen and Chris Potts, 180–193. Somerville, MA: Cascadilla Press.

Ishihara, Shinichiro. 2003. Intonation and interface conditions. Doctoral Dissertation, Massachusetts Institute of Technology.

Ishihara, Shinichiro. 2004. Where does '?' come from?: Syntax-phonology cyclicity mismatch. Talk presented at the Workshop on Prosody, Syntax and Information Structure (WPSI), May 27–29 2004.

Kenstowicz, Michael, ed. 2001. *Ken Hale: A life in language*. Cambridge, MA: MIT Press.

Kitagawa, Yoshihisa, and Janet Dean Fodor. 2003. Default prosody explains neglected syntactic analysis in Japanese. In *Japanese/Korean Linguistics*, ed. William McClure, volume 12, 267–279. Stanford: CSLI Publications.

Kitagawa, Yoshihisa, and Satoshi Tomioka. 2003. Masked island effects in Japanese. Talk presented at Workshop on Altaic Formal Linguistics (WAFL), May 16–18 2003.

Kubozono, Haruo. 1993. *The organization of Japanese prosody*. Tokyo: Kurosio Publishers.

Kubozono, Haruo. 2004. Downstep and phonological phrasing in Japanese. Talk presented at the Workshop on Prosody, Syntax and Information Structure (WPSI), May 27–29 2004.

Kuroda, S.-Y. 1965. Generative grammatical studies in the Japanese language. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Kuroda, S.-Y. 2004. Prosody and the syntax of indeterminates. Talk presented at the Workshop on Prosody, Syntax and Information Structure (WPSI), May 27–29 2004.

Maekawa, Kikuo. 1991a. Perception of intonation characteristics of WH and non-WH questions in Tokyo Japanese. In *Proceedings of the XXIInd International Congress of Phonetic Science*, volume 4, 202–205. Ain-en-Provence, France: Université de Provence.

Maekawa, Kikuo. 1991b. Tokyo hoogen gimon-bun no intoneesyon [The intonation of question sentences in Tokyo dialect]. In *Heisei 3-nendo Nihon Onsei Gakkai Zenkoku Taikai Kenkyuu Happyoo Ronshyuu [Proceedings of the 2001 PSJ Meeting]*, 42–47. Nihon Onsei Gakkai.

Mahajan, Anoop. 1994. Toward a unified theory of scrambling. In *Studies on scrambling: Movement and non-movement approach to free word-order*

phenomena, ed. Norbert Corver and Henk van Riemsdijk, 301–330. Berlin and New York: Mouton de Gruyter.

Miyagawa, Shigeru. 2001. The EPP, scrambling, and *Wh*-in-situ. In Kenstowicz (2001), 293–338.

Nagahara, Hiroyuki. 1994. Phonological phrasing in Japanese. Doctoral Dissertation, University of California, Los Angeles.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer Accademic Publishers.

Pierrehumbert, Janet, and Mary Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A′-movement. In *Alternative conceptions of phrase structure*, ed. M. Baltin and A. Krock, 192–200. University of Chicago Press.

Selkirk, Elisabeth. 1993. The prosodic structure of function words. In *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. James L. Morgan and Katherine Demuth, 187–213. Mahwah, NJ: Lawrence Erlbaum Associates.

Selkirk, Elisabeth. 2000. The interaction of constraints on prosodic phrasing. In *Prosody: Theory and experiment*, ed. Merle Horne, 231–261. Dordrecht and Boston and London: Kluwer Accademic Publishers.

Selkirk, Elisabeth. 2003. Bengali intonation revisited: An optimality theoretic analysis in which FOCUS stress prominence drives FOCUS phrasing. In *Papers in optiomality theory II*, ed. A. Carpenter, A. Coetzee, and P. de Lacy, volume 26 of *University of Massachusetts Occasional Papers (UMOP)*. Amherst, MA: GSLA, University of Massachusetts, Amherst.

Selkirk, Elisabeth, and Koichi Tateishi. 1991. Syntax and downstep in Japanese. In *Interdisciplinary approaches to language: Essays in hornor of S.-Y. Kuroda*, ed. Carol Georgopoulos and Roberta Ishihara, 519–543. Dordrecht: Kluwer Academic Publishers.

Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Shimoyama, Junko. 2001. Wh-constructions in Japanese. Doctoral Dissertation, University of Massachusetts, Amherst.

Sugahara, Mariko. 2003. Downtrends and post-FOCUS intonation in Japanese. Doctoral Dissertation, University of Massachusetts, Amherst, Amherst, MA.

Tomioka, Satoshi. 1997. Wh-insitu, subjacency, and LF syntax. Handout from LSA Summer Institute, Cornel Univesity, Ithaca, NY.

Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and prominence. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Truckenbrodt, Hubert, and Caroline Féry. 2003. More on hierarchical organization and tonal scaling. Unpublished ms.

Yoshida, Keiko, and Tomoyuki Yoshida. 1996. Question marker drop in Japanese. *ICU Language Research Bulletin* 11:37–57.

Yoshida, Tomoyuki. 1998. *Wh*-operator vs. *Yes/No*-operator. *ICU Language Research Bulletin* 13:159–172.

Shinichiro Ishihara
Universität Potsdam
SFB632 „Informationsstruktur"
Postfach 601553
14415 Potsdam
Deutschland
ishihara@rz.uni-potsdam.de
http://alum.mit.edu/www/s_i/

The Influence of Tense in Adverbial Quantification*

Cornelia Endriss and Stefan Hinterwimmer†*

*Universität Potsdam and †Humboldt-Universität zu Berlin

We argue that there is a crucial difference between determiner and adverbial quantification. Following Herburger [2000] and von Fintel [1994], we assume that determiner quantifiers quantify over individuals and adverbial quantifiers over eventualities. While it is usually assumed that the semantics of sentences with determiner quantifiers and those with adverbial quantifiers basically come out the same, we will show by way of new data that quantification over events is more restricted than quantification over individuals. This is because eventualities in contrast to individuals have to be located in time which is done using contextual information according to a pragmatic resolution strategy. If the contextual information and the tense information given in the respective sentence contradict each other, the sentence is uninterpretable. We conclude that this is the reason why in these cases adverbial quantification, i.e. quantification over eventualities, is impossible whereas quantification over individuals is fine.

Keywords: Adverbial Quantification, Covert Variables, Domain Restriction, Quantificational Variability

1   Data

It is usually assumed (cf. Lewis [1975], Heim [1982], von Fintel [1994], Chierchia [1995], Kratzer [1995], Herburger [2000] and many others) that the interpretation of A(dverbially)-quantified sentences such as (1-a) comes out the

same as the interpretation of (1-b) with a D(eterminer)-quantifier. This is generally referred to as a quantificational variability (QV) effect.

(1)   a.   A police car is usually green.

       b.   Most police cars are green.

But whereas this is true for the above example, it does not hold in general that A-quantified sentences and the corresponding D-quantified ones end up with the same interpretation. We will present data which show that there are A-quantified constructions which are generally judged to be uninterpretable (cf. (2)), though the D-quantified versions of them are considered perfectly fine (cf. (3)):

(2) $^{??}$A car that was bought in the 80s is usually blue.

(3)   Most cars that were bought in the 80s are blue.

The indefinite DP in (2) only seems to get a specific reading with scope over the Q-adverb. This interpretation results in a deviant reading, as the property of having some specific colour is stable for a given car under normal circumstances, i.e. the predicate *to be blue* is usually interpreted as an individual level predicate with respect to cars[1]. This raises the question of why the reading where the Q-adverb has scope over the indefinite DP is blocked in (2).

   Interestingly, (4) is much better than (2):

(4)   A car that was bought in the 80s was usually blue.

In contrast to this, (5) is just as good as (3) though different in interpretation[2]:

(5)   Most cars that were bought in the 80s were blue.

---

[1] Of course, cars can change their colour when they are painted differently, which means that, strictly speaking, *blue* is not an individual level predicate in this context. Yet, we will stick to this assumption in the following.

[2] We will discuss the interpretative difference in section 4.2.

2 Existing Analyses

In this section we want to show that existing analyses cannot explain the difference in acceptability between (2) and (3).

2.1 Q-adverbs as unselective binders

In the theories of Heim [1982] (whose theory is based on Lewis [1975]), Kamp [1981], Diesing [1990], and Kratzer [1995], indefinites provide a restricted variable. If the sentence does not contain a Q-adverb, the restricted variable is subject to existential closure. Otherwise, it is bound by an adverbial quantifier. Adverbial quantifiers are unselective binders that bind every free variable in their scope, i.e. individual as well as situation/event variables. Stage level predicates come with a spatio-temporal argument whereas individual level predicates do not.

Despite its oddity, (2) gets a perfectly coherent interpretation according to these approaches, as there is a free variable (provided by *a car*) which can be bound by the quantifier *MOST*, which is the denotation of *usually*.

(6)   a. $^{??}$A car that was bought in the 80s is usually blue.
        b. $MOST_x \left[car(x) \wedge bought\_in\_80s(x) \right] \left[blue(x) \right]$

This is exactly the same interpretation as the ones that is assigned to (3):

(7)   a.   Most cars that were bought in the 80s are blue.
        b. $MOST_x \left[car(x) \wedge bought\_in\_80s(x) \right] \left[blue(x) \right]$

This means that these theories cannot adequately account for the acceptability differences.

2.2   Q-adverbs as topic sensitive binders

Chierchia [1995] differs from the above view in two respects: Firstly, indefinites are interpreted as regular existential quantifiers. When they are topic marked, they are existentially disclosed and can be bound by a c-commanding adverbial quantifier afterwards. And secondly, individual level predicates also come with a spatio-temporal argument, but in contrast to a stage level argument it needs to be bound by the generic quantifier.

(8)   a. $^{??}$A car that was bought in the 80s is usually blue.

b. $MOST_x \big[car(x) \wedge bought\_in\_80s(x) \big]$
$$\big[GEN_s \big(x \; in \; s \big) \big(blue(x, s) \big) \big]$$

This interpretation is equivalent to the interpretation of (3) (as shown in (7)); and here again, there is no reason why the sentence should be unacceptable.

2.3   Situation and event semantic approaches

Following de Swart [1993], von Fintel [1994], and Herburger [2000], Q-adverbs bind situation or event variables. Indefinites are ordinary existentially quantified DPs. Quantificational variability then results from binding (minimal) situations/events that contain just one individual of the relevant sort. It is important that for each situation, a different individual is chosen so that the the individuals vary with the situations (cf. von Fintel [1994]). This in turn guarantees the quantificational variability effect. The restriction and the nuclear scope of the respective Q-adverb are determined on grounds of information structure or contextual information.

Even in these theories, the semantic representation of (2) still comes out equivalent to the semantics of (3) shown in (7). This means that without further assumptions, the situation/event semantic accounts also cannot explain the ob-

served acceptability differences, as there is a perfectly coherent representation for (2) in these approaches:

(9)　　a.　$^{??}$A car that was bought in the 80s is usually blue.

　　　b.　$MOST_e \left[\exists x.arg(e,x) \land car(x) \land [\exists e'.buy(e') \land theme(e',x) \right.$
　　　　　　　　　　　　　　　　$\left. \land in\_the\_80s(e')]\right] \left[blue(e)\right]$

3　Conceivable Solution Strategies

As has been shown in the preceding section, none of the existing theories can explain the difference in acceptability between (2) and (3). Before we will present our account of these data, we want to mention briefly two alternative solution strategies that could come to mind, and argue why they cannot be maintained.

3.1　Natural classes?

One could speculate that QV is only possible with indefinites that pick out individuals from a well defined class (cf. Krifka et al. [1995] and Cohen [2001] for generics and natural classes; Greenberg [2002] and Greenberg [2003] for the different behaviour of singular indefinites and bare plurals in generic sentences, i.e. sentences that do not contain an overt Q-adverb[3]). But the fact that the following sentence is perfectly acceptable shows that this cannot be the correct generalization for the cases discussed here:

(10)　　A French linguist with green hair and six toes is usually intelligent.

It will be hard to argue that the class of *French linguists with green hair and six toes* is a natural one or even that this should be a more natural class than the one

---

[3]　We would like to thank Angelika Kratzer, who drew our attention to the work of Yael Greenberg.

of cars that have been bought in the 80s referred to in the initial example (2).

There also seems to be a difference between temporal and spatial restrictions. Whereas the restriction of the existentially bound variable by a property that refers to a specific time interval renders a sentence with an individual level predicate ungrammatical (as in (2)), restricting it by a property that refers to a specific location is harmless:

(11)   A car that is bought in the car store in Fleet Street is usually blue.

3.2   Specificity?

Alternatively, it could be argued that for some unknown, yet compelling reason, temporally fixed indefinites have to be interpreted specifically. But this assumption is also not borne out as the generalization does not hold for non-QV environments:

(12)   It is possible that a car that was bought in the 80s may have had an accident today.

(13)   Every customer recognized a car that was on exhibition in this shop window yesterday.

In (12), the speaker does not need to have a particular car in mind, and in (13) the cars may vary with the customers. This shows that the reason for the unacceptability of (2) cannot be a forced specific interpretation for the indefinite.

4   A Pragmatic Account

We follow von Fintel [1994] and Herburger [2000] in the assumption that D-quantifiers take sets of individuals as arguments, while A-quantifiers take sets of eventualities. The arguments of D-quantifiers are determined grammatically,

while the restriction of A-quantifiers has to be determined solely on the basis of information structure (or contextual information).

We also assume that every quantification comes with a domain restriction (cf. von Fintel [1994], Martí [2003], Stanley [2000], and Stanley [2002]). For individual quantifiers this means that the restrictor set has to be intersected with the denotation of a covert predicate that is determined by the context. In a context[4] as given in (14-a), a sentence such as (14-b) would not be about all apples of the universe, but about all the apples that have been introduced in the previous sentence, i.e. all apples that Peter bought the day before:

(14)   a.   Yesterday, Peter bought apples.

            b.   Every apple tasted awful.

Analogously, domain restriction for events means locating the respective events in time (cf. Partee [1973], Lenci and Bertinetto [1999]). In a context such as (i-a), the event of drinking beer in (i-b) is interpreted as taking place at the same time as the contextually given eventuality in (i-a), i.e. during the time when Peter was at Mary's party (cf. Partee [1973]):

(15)   a.   Yesterday, Peter had a good time at Mary's party.

            b.   He drank beer.

We now claim that the acceptability differences between the initial examples (2) and (3) can be explained on the basis of (conflicting) tense information.

---

[4]   Also extra-linguistic contexts can serve to restrict the quantifier domain as in the well-known example of Lewis [1986]:

(i)   a.   *When looking into the fridge, someone says:*
       b.   There is no beer.

Here, (i-b) would not be about beer in general, but only about beer in the respective fridge due to the contextual situation given in (i-a).

4.1   Technical preliminaries

We will explain the technical apparatus by first looking at (3) and explaining why this is a good sentence for which there exists a sensible interpretation that is predicted by our approach.

Due to the presence of the D-quantifier *most*, the sentence is interpreted by employing quantification over individuals x. As every quantification comes with a domain restriction, so does the quantifier *most*, and an additional conjunct $C(x)$ is introduced.[5] Every verbal predicate introduces a variable, which in the default case is bound by an existential quantifier. This quantifier also comes with a domain restriction.

(16)   a. Most cars that were bought in the 80s are blue.

b. $\underline{MOST_x}\ \big[car(x) \wedge [\underline{\underline{\exists e'}}.buy(e') \wedge theme(e', x) \wedge past(e')$
$$\wedge in\_the\_80s(e') \wedge \underline{C(e')}] \wedge \underline{C'(x)}\big]$$
$$\big[\underline{\underline{\exists e}}.arg(e, x) \wedge pres(e) \wedge blue(e) \wedge \underline{C(e)}\big]$$

To indicate that a context restriction belongs to a quantifier, we have underlined the corresponding terms in formula (16). In case of quantification over eventualities, the restriction temporally locates an eventuality e within an interval i_e. This means that C is of the form $e\ @\ i_e$:

(17)   a. Most cars that were bought in the 80s are blue.

b. $MOST_x\ \big[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge past(e')$
$$\wedge in\_the\_80s(e') \wedge e'\ @\ i_{e'}] \wedge C'(x)\big]$$
$$\big[\exists e.arg(e, x) \wedge pres(e) \wedge blue(e) \wedge e\ @\ i_e\big]$$

Temporal location of an event within an interval is defined as follows:

---

[5]   Note that in contrast to von Fintel [1994] and Martí [2003] we assume that this domain restriction is added at the latest possible position, because it is determined by overt information that has been mentioned before.

(18) $\qquad e \ @ \ i_e := \tau(e) \subseteq i_e,$

$\qquad$ *where $\tau(e)$ denotes the running time of e.*

In words, $e \ @ \ i_e$ means that e—in case of verbs denoting dynamic eventualities (i.e. achievements, accomplishments and activities, see Vendler [1957])—takes place /holds at some time during the interval i_e or, in case of a stative verb/property, exhausts i_e[6].

We assume the following (simplified) semantics for tense information relative to the speech time t_0:

(19) $\qquad$ a. $pres(e) := t_0 \in \tau(e)$
$\qquad\qquad$ b. $past(e) := \tau(e) < t_0$

4.2 The interval resolution strategy

The free interval variables i in (17) have to be fixed by overt or contextual information as far as it is available.

We assume the following pragmatic strategy for the temporal localization of the events, what we will call the **interval resolution strategy**:

(20) $\qquad$ 1.Take overt information.

$\qquad$ 2.If not available: Take contextual information from the same domain (restrictor vs. nucleus), i.e. the running time of another salient

---

[6] Following Bach [1986] (among many others, see also Rothstein [2003] and references therein for a recent discussion), we assume that statives (as well as activities) are homogenous with respect to their internal structure. In case of stative verbs such as *to be French*, the state of *being French* for a given individual denotes an infinite set of *being French* eventualities the largest of which is the maximal eventuality in which the property of being French holds for the individual under consideration. Under this view, it follows that $e \ @ \ i_e$ picks out those subeventualities of the state under discussion that lie in the interval i_e. Analogous to the case of activities, only the maximal eventuality (i.e. the one exhausting the whole interval) is taken into account when computing the truth conditions of the sentence. This is because quantification over infinite sets is no reasonable operation.

event.

3. If not available: Take contextual information from the other domain. Or take the default time interval i_{world}, which denotes the whole time axis.

The principle behind this strategy is the following: If there is overt information about the time in which a respective event e has to be located, this information has to be taken to instantiate the interval i_e. This would be the case in example (15-a) where the event of Peter's having a good time at Mary's party has to be located during the interval denoted by *yesterday*. In (15-b) on the other hand, there is no overt interval in which the beer drinking event has to be located. Here, contextual information has to be taken into account—which corresponds to point (2.) of the interval resolution strategy given in (20). According to this strategy, the event of Peter's beer drinking is interpreted during some contextually given time interval which in this case is the running time of some other contextually given salient event, i.e. the time when Peter was at Mary's party. The concept of local proximity plays a role here. Contextual information which has been introduced immediately before the event to be located is more appropriate to function as restriction for this event than material that has been presented much earlier. This is reflected in the interval resolution strategy in (20) where local information (point 2.) is to be preferred over non local one (point 3.).

Quantification over individuals

In case of (17), repeated here as (21), there are two intervals which have to be resolved: i_e and $i_{e'}$.

(21)   a. Most cars that were bought in the 80s are blue.
        b. $MOST_x \left[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge past(e') \right.$

$$\wedge in\_the\_80s(e') \wedge e' @ \underline{i_{e'}}] \wedge C'(x)\big]$$
$$\big[\exists e.arg(e, x) \wedge pres(e) \wedge blue(e) \wedge e @ \underline{i_e}\big]$$

For the relative clause event e' which has to be located in the interval $i_{e'}$, there is overt information, i.e. *the 80s*. The interval has to be instantiated with the explicitly mentioned interval *the 80s*. Concerning i_e, there is neither overt information in the matrix clause nor any other interval information in the same domain, which is the nucleus. Therefore the third option of the interval resolution strategy in (20) comes into play and the interval could be resolved contextually by taking information from the other domain, i.e. by the running time of e':

(22)   a. Most cars that were bought in the 80s are blue.
         b. $\#MOST_x \big[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge past(e')$
$$\wedge in\_the\_80s(e') \wedge e' @ \underline{80s}] \wedge C(x)\big]$$
$$\big[\exists e.arg(e, x) \wedge pres(e) \wedge blue(e) \wedge e @ \underline{\tau(e')}\big]$$

The event e would then be interpreted as being located within the same period as e', which is during the 80s. But this would directly clash with the semantics of present tense:

(23)   a. Most cars that were bought in the 80s are blue.
         b. $\#MOST_x \big[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge \tau(e') < t_0$
$$\wedge in\_the\_80s(e') \wedge \tau(e') \subseteq 80s] \wedge C(x)\big]$$
$$\big[\exists e.arg(e, x) \wedge \underline{t_0 \in \tau(e)} \wedge blue(e) \wedge \underline{\tau(e) \subseteq \tau(e')}\big]$$

Formula (23) is inconsistent with the situation that the speech time t_0 is not contained in the eighties:

$$t_0 \notin \tau(e) \subseteq \tau(e') \subseteq 80s$$

The only option left for the interval resolution strategy to create a coherent interpretation is to instantiate the time interval with the whole time axis:

(24)   a. Most cars that were bought in the 80s are blue.

b. $MOST_x \; \big[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge \tau(e') < t_0$
$\wedge in\_the\_80s(e') \wedge e' \; @ \; 80s] \wedge C(x)\big]$
$\big[\exists e.arg(e, x) \wedge t_0 \in \tau(e) \wedge blue(e) \wedge e \; @ \; \underline{i_{world}}\big]$

This then means: *Most cars bought in the 80s are presently blue.*

However, in (5)—the variant of (2), in which the matrix predicate is set to past tense—the interval of the matrix clause can be set to the running time of e'. Here, there is no time clash due to the past tense marking of the matrix clause verb:

(25)   a. Most cars that were bought in the 80s were blue.

b. $MOST_x \; \big[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge past(e')$
$\wedge in\_the\_80s(e') \wedge e' \; @ \; 80s] \wedge C(x)\big]$
$\big[\exists e.arg(e, x) \wedge \underline{past(e)} \wedge blue(e) \wedge \underline{e \; @ \; \tau(e')}\big]$

The meaning is: *Most cars bought in the 80s were blue when they were bought.*

But still it would be possible to set the interval to the whole time axis according to point (3.) of the strategy above. This leads to a different reading for this sentence that indeed seems to be available:

(26)   a. Most cars that were bought in the 80s were blue.

b. $MOST_x \; \big[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge past(e')$
$\wedge in\_the\_80s(e') \wedge e' \; @ \; 80s] \wedge C(x)\big]$
$\big[\exists e.arg(e, x) \wedge past(e) \wedge blue(e) \wedge e \; @ \; \underline{i_{world}}\big]$

The past tense demands $\tau(e)$, i.e. the time of being blue, to end before the speech time t_0:

(27)   a. Most cars that were bought in the 80s were blue.

b. $MOST_x \left[car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x) \wedge \tau(e') < t_0 \right.$
$\left. \wedge in\_the\_80s(e') \wedge e' @ 80s] \wedge C(x) \right]$
$\left[\exists e.arg(e, x) \wedge \underline{\tau(e) < t_0} \wedge blue(e) \wedge e @ i_{world} \right]$

This means that the *blue* eventuality has to end before the speech time. Under the assumption that *blue* is regarded as an individual level predicate with respect to cars, this triggers the expectation on the side of the hearer that the respective cars do not exist any longer. We take this to be a consequence of our analysis of individual level predicates: On the one hand, only the maximal eventualities of cars being blue that lie within the respective interval which is i_{world} are picked out (see footnote 6). On the other hand, past tense marking would keep it from doing so if those cars would still exist (without having changed their colour). This is because past tense forces those eventualities to end before the speech time, while there are larger eventualities of the cars being blue that lie within the interval i_{world}: namely those comprising the whole time of existence of the cars. That means, using past tense one would not give as much information with respect to the chosen interval as possible, if the cars would still exist. If, on the other hand, the cars do not exist anymore, past tense marking would allow to pick out the largest eventualities of the respective cars being blue that lie within the given interval i_{world}. Therefore, the hearer automatically assumes that the respective cars indeed do not exist anymore[7].

---

[7]   As has been pointed out to us by Manfred Krifka, there is another possibility for resolving i_e in case of (5) (repeated here as (i-b)), namely to a contextually salient interval. In a context such as (i-a), it would be the year of 1995 or more precisely the time when the second-hand car market took place:

(i)   a.   Talking about the second-hand car market in 1995.

        b.   Most cars that were bought in the 80s were blue.

This is predicted by our approach because according to point (3.) of the interval resolution strategy, non-local contextual information can be taken into account.

This effect is reminiscent of the facts discussed by Kratzer [1995] and Musan [1997] under the label *life time effects*. Consider the sentence in (28):

(28)   Gregory was from America.

If the sentence is uttered out of the blue, it implicates that Gregory is dead at the speech time (or has changed his citizenship). The very same effect arises in the second reading of (5) given in (27).

To summarize the findings of this section, we claim that (3) is fine for the following reasons:

- D-quantification does not bind eventualities.

- The predicate *to be blue* in the nuclear scope introduces an existentially bound eventuality variable e.

- This eventuality is located in an interval that is independent of the one given in the relative clause.

- There is no interval information in the nuclear scope.

- The interval i_e can be set to the default interval i_{world}.

Quantification over eventualities

In case of (2), repeated as (29), matters are different.

(29) $^{??}$A car that was bought in the 80s is usually blue.

Regarding adverbial quantification, it is not the syntax that determines restrictor and nucleus, but information structure (or contextual information): Non-focal/topical material is mapped onto the restrictor, focal material is mapped

onto the nuclear scope (cf. among others Chierchia [1995], Krifka [1995], Partee [1995], Rooth [1995], Herburger [2000]). In this example, the matrix predicate *blue* is focussed, and therefore it is mapped onto the nuclear scope. Furthermore—and this is crucial for our account—the eventuality variable introduced by *blue* is bound by the adverbial quantifier *usually* in the restrictor as well as in the nuclear scope. This has the consequence that the eventuality variable introduced by the matrix verb ends up in the same domain as the eventuality variable introduced by the relative clause internal verb—namely in the restrictor of the adverbial quantifier *usually*. This contrasts with the situation in (3), where the two variables are interpreted in different domains: The variable introduced by the relative clause verb is interpreted in the restrictor of the determiner quantifier *most*, while the variable introduced by the matrix verb ends up in the nuclear scope of this quantifier. This, together with the fact that the matrix eventuality variable also needs to be restricted by a time interval, has the consequence that the interval resolution strategy given in (20) works differently in the two cases. Now consider the semantic representation of (2) (repeated here as (30)) in detail:

(30)   a. $^{??}$A car that was bought in the 80s is usually blue.

b. $MOST_e \left[\exists x . arg(e, x) \wedge car(x) \wedge [\exists e' . buy(e') \wedge theme(e', x) \right.$
$\left. \wedge past(e') \wedge in\_the\_80s(e') \wedge C(e')] \wedge C'(x) \wedge C(e) \right]$
$\left[pres(e) \wedge blue(e) \right]$

As mentioned above, the domain restriction $C(e)$ for the adverbial quantifier *usually* must include the constraint $e @ i_e$, where i_e has to be resolved. As there is no overt information with respect to i_e in the matrix clause, the only available interval information originates from the information about the event e' in the relative clause. This is information originating from the same domain, i.e. from the restrictor, and according to the interval resolution strategy in (20), i_e has to

be equated to the interval denoted by the running time[8] of e':

(31)   a. $^{??}$A car that was bought in the 80s is usually blue.
        b. $MOST_e \ \big[\exists x.arg(e,x) \wedge car(x) \wedge [\exists e'.buy(e') \wedge theme(e',x)$
              $\wedge past(e') \wedge in\_the\_80s(e') \wedge e' \ @ \ 80s] \wedge C'(x) \wedge e \ @ \ \underline{\tau(e')}\big]$
              $\big[pres(e) \wedge blue(e)\big]$

As e' takes place in the 80s and e is located during the running time of e', only events located in the 80s, i.e. before the speech time t_0, will be considered in the restrictor whereas the nucleus requires the events to include the speech time:

(32)   a. $^{??}$A car that was bought in the 80s is usually blue.
        b. $MOST_e \ \big[\exists x.arg(e,x) \wedge car(x) \wedge [\exists e'.buy(e') \wedge theme(e',x)$
              $\wedge \tau(e') < t_0 \wedge in\_the\_80s(e') \wedge \tau(e') \subseteq 80s] \wedge C'(x) \wedge \underline{\tau(e) \subseteq \tau(e')}\big]$
              $\big[\underline{t_0 \in \tau(e)} \wedge blue(e)\big]$

This necessarily yields an empty intersection of restrictor and nucleus and thus accounts for the oddity of (2).

As this oddity is not due to a grammatic but due to a pragmatic principle, it is to be expected that the unacceptability is not absolute. For some speakers it might be possible to construct contexts in which the sentence is fine for them. Still, (2) will be much less natural than (3), where it is not necessary for the speaker to construct a matching context to be able to interpret the sentence adequately.

Obviously, if the information in the matrix clause is non-contradictory in this respect, one expects the utterance to be felicitous, which is in fact borne

---

[8] Compare this to example (15-b), where, in the given context, the second event of Peter's beer drinking has to be interpreted in the running time of the eventuality of the first sentence—i.e. when he was at Mary's party—due to the local proximity of the two sentences. We assume that the mechanism is the same in the case discussed in (30). Here also, one cannot help but interpret the sentence with the interval i_e set to the running time of the the salient relative clause event, as this is local information.

out. This can be seen in (4), repeated here as (33):

(33)   a. A car that was bought in the 80s was usually blue.
        b. $MOST_e \, \big[\exists x.arg(e, x) \wedge car(x) \wedge [\exists e'.buy(e') \wedge theme(e', x)$
              $\wedge past(e') \wedge in\_the\_80s(e') \wedge e' \text{ @ } 80s] \wedge C(x) \wedge \underline{e \text{ @ } \tau(e')} \big]$
              $\big[\underline{past(e)} \wedge blue(e) \big]$

Here, instantiating i_e with the running time of e' which is located in the 80s does not lead to a contradiction with the past tense information in the nucleus.

Our approach predicts example (2) to be out for the following reasons:

- A-quantification binds the eventuality variable e in the restrictor **and** in the nuclear scope.

- Domain restriction forces e to be located in an interval i_e.

- Due to contextual information in the restrictor, i_e has to be resolved to the running time of e', which is located in the past.

- This clashes with the present tense information in the nuclear scope.

- The intersection of restrictor and nucleus is necessarily empty.

4.3   Explicit interval setting

Consider (34), which is fine in spite of the fact that it is structurally almost identical to (2): The matrix verb is marked for present tense, while the verb in the relative clause is marked for past tense. Obviously, what makes the difference is the presence of the adverb *nowadays* in the matrix clause[9].

(34)   a. A car that was bought in the 80s is usually rusty nowadays.
        b. $MOST_e \, \big[\exists x.arg(e, x) \wedge car(x) \wedge nowadays(e) \wedge [\exists e'.buy(e')$

---

[9]   We assume that *nowadays* is not focussed and hence mapped onto the restrictor.

$$\wedge theme(e', x) \wedge past(e') \wedge in\_the\_80s(e') \wedge e' @ 80s]$$
$$\wedge C(x) \wedge e @ i_e\big]\ \big[pres(e) \wedge rusty(e)\big]$$

Let us assume for concreteness that *nowadays* introduces an interval of contextually specified size which is constrained to include the speech time, and locates the eventuality introduced by the verb it modifies within this interval.[10] As this is overt information, (34) is predicted to be fine by the interval resolution strategy given in (20): The interval i_e does not need to be set to the running time of the eventuality denoted by the relative clause verb, but can or—according to point (1.) of the interval resolution strategy given in (20)—has to be set to the interval denoted by the overt interval information introduced by *nowadays*. In this case, there is no clash between the temporal information in the restrictor and the temporal information the present tense marking of the matrix verb contributes to the nuclear scope:

(35)   a. A car that was bought in the 80s is usually rusty nowadays.

   b. $MOST_e\ \big[\exists x.arg(e, x) \wedge car(x) \wedge nowadays(e) \wedge [\exists e'.buy(e')$
$$\wedge theme(e', x) \wedge past(e') \wedge in\_the\_80s(e') \wedge e' @ 80s]$$
$$\wedge C(x) \wedge \underline{e @ nowadays}\big]\ \big[\underline{pres(e)} \wedge rusty(e)\big]$$

As can easily be seen, the present tense information in the matrix clause does not clash with the interval information of the restrictor, and the sentence is therefore felicitous.

---

[10] As has been pointed out to us by Manfred Krifka and Alex Grosu, it is not obvious why *nowadays* introduces such an interval whereas present tense does not and therefore does not lead to an interval resetting. Possibly, *nowadays* behaves just like *still* and *meanwhile* in that it presupposes an interval in the past (cf. the following two subsections), which would be an alternative explanation for the felicity of (34). We will have to leave this question for future work.

4.4   Interval setting induced by presuppositions

Just as (34), also (36) is fine, in spite of differing tenses in matrix and relative clause. This seems to be due to the presence of the adverbial *still* in the matrix clause.

(36)   A car that was bought in the 80s is usually still roadworthy.

First, we assume that *still* is similar to *nowadays* in that it introduces an interval in which the event e has to be located. Besides that, it does not add much to the semantic content:

(37) $\text{still}(P, e) = P(e) \wedge e \ @ \ t$, *where e is the eventuality of the matrix event predicate P (be roadworthy).*

We assume that *still* takes two arguments: As a first argument, it takes the eventuality predicate P denoted by the intermediate verbal projection that it modifies and that has already been applied to all its individual arguments. We assume these arguments to be base generated inside the verbal projection (cf. Koopman and Sportiche [1991]). Therefore, the eventuality predicate P denotes a function from eventualities to truth values. The second argument is the eventuality variable introduced by the respective verb. In line with Kratzer [1995], we assume that the eventuality arguments of verbs are directly represented in the syntax: They are generated in the outermost specifier position of the verbal projection. Under the assumption that *still* is adjoined directly below the eventuality argument, it first combines with the denotation of the intermediate verbal projection below it, and in the next step combines with the respective eventuality variable.

What is crucial for our purposes is that apart from its rather trivial assertion, *still* also triggers a presupposition (cf. Löbner [1999], Smessaert and ter Meulen [2004], among others):

(38) $\exists t'.salient(t') \wedge t' < t \wedge \forall t''.[t' \leq t'' < t \rightarrow \exists e'.e' @ t'' \wedge P(e')]$,
where t is the time interval that is introduced by its lexical content, cf. (37).

For this presupposition to be satisfied in the case of (36), there has to be a salient time interval t' which is located before t where the eventuality e of being roadworthy held. This property has to persist during the time until t starts. In this example, the explicitly mentioned interval denoted by the eighties can serve to locally satisfy the presupposition: It is plausible to assume that the respective cars already had the property of being roadworthy at the time when they were bought.

As before, the overtly introduced interval t (originating from the lexical content of *still*) serves to determine the interval i_e. As t follows t', which is set to the 80s due to the presupposition binding, t is an interval following the 80s and can thus include the speech time.

(39)   a. A car that was bought in the 80s is usually still roadworthy.
        b. $MOST_e \left[\exists x.arg(e,x) \wedge car(x) \wedge [\exists e'.buy(e') \wedge theme(e',x) \right.$
                  $\left. \wedge past(e') \wedge in\_the\_80s(e') \wedge e' @ 80s] \wedge C(x) \wedge e @ t \right]$
           $\left[pres(e) \wedge roadworthy(e) \right]$,
where t follows the 80s due to presupposition resolution.

Basically the same reasoning applies to the following example[11]:

(40)   A car that was bought in the 80s usually broke meanwhile.

---

[11] The sentences in (34), (36), and (40) are construed as parallel as possible to the initial example sentence (2). But as the respective sentences cannot reasonably be uttered with individual level predicates (which *blue* is assumed to be with respect to cars), the matrix predicate had to be substituted. As can be seen in the following, the sentences are out with true individual level predicates:

We assume that *meanwhile* has the same lexical content as *still*, but introduces a different presupposition:

(41) $\exists t'.salient(t') \wedge t' < t \wedge [\neg\exists e'.e' @ t' \wedge P(e')]$

As it is plausible to assume that the respective cars did not have the property of having been broken at the time when they were bought, (40) is also predicted to be fine: Again, the presupposition introduced by the adverb can be satisfied locally, and the matrix eventualities can be located in an interval that is compatible with the present tense information in the nuclear scope.

5   Causally Related Eventualities

The following examples are all fine, in spite of the fact that each of them exemplifies the constellation that led to pragmatic deviance in our initial set of examples, i.e. the relative clause verbs are marked for past tense, while the matrix verbs are marked for present tense, and there is no overt interval setting:

(42)   A car that was made in the 80s is usually blue.

(43)   A house that was built in the 19th century usually has a gabled roof.

(44)   A lawyer who was educated in Berlin is usually competent.

(45)   A man who was in jail during the 80s usually has a Bruce Lee tatoo.

(i) $^{??}$A car that was bought in the 80s is usually still a BMW.

(ii) $^{??}$A car that was bought in the 80s is usually a BMW meanwhile.

In the case of *meanwhile*, the presupposition can never be fulfilled whereas in case of *still*, the temporal adverbial is superfluous as it only adds a presupposition which is already introduced by the very definition of an individual level predicate.

What all the sentences have in common is that the states denoted by the matrix verbs are interpreted as being (at least indirectly) caused by the relative clause eventualities. In examples (42) to (44), the relative clause internal predicate denotes a set of telic events. The sentences are interpreted as saying that the culmination point of the telic event coincides with the matrix state. With verbs of creation as the ones given in (42) and (43), this is trivially true, because properties are usually only ascribed to existing entities. In (44), this is due to the specific relation between the relative clause event and the matrix state. In (45), where the relative clause internal predicate denotes a state without a culmination point, it is still required that the matrix state does not hold of the respective individual when the relative clause internal eventuality starts.

Once a different predicate is chosen in the matrix clause, the sentences become odd. Compare (44) to (46):

(46) $^{??}$A lawyer who was educated in Berlin is usually blond.

The reason for the felicity of examples (42) to (45) seems to be the fact that it is impossible to convey the correct meanings of the sentences by using past tense in both relative and matrix clause. To put it differently, the possibility of expressing the correct meaning of the respective sentence with past tense as in (4) blocks the possibility to use present tense for the matrix clause (as in (2)).

Consider an example similar to (43), but with past tense also in the matrix clause:

(47)   A house that was built in the 19th century usually had a gabled roof.

This sentence either means that at least some houses with the respective property do not exist any more at the speech time (which is a reading with a life time effect as described for (5)) or that houses that were built in the 19th century used to have a gabled roof before they were built, which is a very implausible

reading.

According to the interval resolution strategy given in (20), this is predicted. If e (where e is the eventuality of having a gabled roof) is interpreted as holding at the same time as e' (where e' denotes the time when the relative clause internal event takes place), the corresponding representation for (47) is as follows:

(48) $MOST_e \left[\exists x.arg(e, x) \wedge house(x) \wedge [\exists e'.build(e') \wedge theme(e', x) \right.$
$\wedge\, past(e') \wedge 19c(e') \wedge e' @ 19c] \wedge C(x) \wedge e @ \tau(e')]$
$\left[past(e) \wedge gabled\_roof(e)\right]$

This would imply that the gabled roof was already a property of the respective houses when they were built. This is not what sentence (43) is supposed to express.

If, on the other hand, the third step of the interval resolution strategy in (20) is taken, and the matrix interval is set to the whole time axis, the sentence comes to mean that most (maximal) eventualities that stand in a thematic relation to a house that was built in the 19th century are eventualities of having a gabled roof that end before the speech time. This however implies that the respective houses do not exist anymore, and a life time effect obtains. This does not correspond to the intended meaning of (43) either. Furthermore, it means violating the interval resolution strategy given in (20), as this would only allow the matrix interval to be set to the running times of the respective relative clause eventualities.

Therefore the strategy which was helpful before (example (4)), namely to set the matrix predicate to past tense, is no way to go in the above examples. In that case, according to the interval resolution strategy given in (20), i_e would be instantiated with the interval that denotes the running time of the relative clause eventuality. But here, e does not hold at this stage. So the only way to express the intended meaning of the sentence is to set the matrix predicate to present tense and then directly take the last step of the interval resolution strategy given

in (20) and instantiate the interval i_e with the whole time axis[12].

The proposed mechanism seems to be confirmed by the following facts:

(49)   A lawyer who was educated in Berlin was usually competent.

(50)   A man who was in jail during the 80s usually had a Bruce-Lee tatoo.

In (49) and (50), either a life time effect is triggered or in case of (49), the sentence is interpreted in a way that the state of being competent was already true at the time the education event started. In case of (50), the sentence gets the interpretation that the state of having a Bruce Lee tatoo was already true for a person before the respective person came to jail. This is predicted because the relative clause internal event and the matrix predicate are assumed to take place at the same time according to the interval resolution strategy.

Therefore, there is no other possibility to express the intended meaning than to use present tense in the respective matrix clauses. This accounts for the felicity of (42) to (45)[13].

---

[12] Point (1.) of the interval resolution strategy cannot be applied, because there is no overt information. Point (2.) is no option either as this would lead to the same contradiction as shown for example (2).

[13] As Graham Katz has pointed out to us, there are related data which are problematic for our account:

(i)   A song that was popular in the 80s usually has electronic beats in it.

Though it is not only possible, but necessary that the respective songs already had electronic beats in them when they were popular, the sentence is still perfectly fine. We can only speculate that this could be due to the fact that here also, the intended meaning of the sentence cannot adequately be expressed by the past tense variant of it. This could be because only present tense in the matrix clause expresses that the songs still exist at the speech time.

6   Summary

Based on a set of new observations, we have argued for an analysis of Q-adverbs as (exclusive) binders of eventuality variables. We have shown that the availability of QV-readings in sentences with indefinite DPs containing a relative clause is sensitive to the interaction of the tense markings of the respective clauses (matrix clause vs. relative clause): In the standard case, QV is only possible if the tenses agree. We have argued for the existence of a pragmatic strategy that temporally locates the eventualities bound by the Q-adverb in an interval that is determined on the basis of available information. This pragmatic mechanism is sensitive to locality considerations: In the absence of overt information, it locates the eventualities quantified over in the same interval as the running times of the respective relative clause eventualities, as these count as interval information originating from the same domain (i.e. the restrictor). If this information about the temporal location of the respective eventualities contradicts the information constituted by the tense marking of the respective matrix verbs (which are interpreted in the nuclear scope), the resulting structures are semantically deviant. We have explained why in certain well defined cases the interval resolution strategy given in (20) does not rule out the otherwise infelicitous structures from above. This is either due to the presence of adverbs that overtly introduce an interval in which the eventualities can be located, or due to a specific relation holding between the relative clause and the matrix eventualities: If the matrix eventualities can naturally be interpreted as having been (at least indirectly) caused by the relative clause eventualities, the respective sentences are fine. We accounted for this effect by showing that skipping an otherwise obligatory step of the interval resolution strategy and resolving the contextual variable responsible for the temporal location of eventualities to the whole time axis is the only way to express the intended meanings of the respective clauses, i.e. to express the (sometimes indirect) causal relations between the respective

eventualities.

7   Outlook

As Alex Grosu has pointed out to us, the grammaticality difference between (51) and (52) seems to have a similar origin as the acceptability differences of the data discussed in this paper:

(51)   *A car that would be designed by Mary is usually blue/will usually be blue.

(52)   A car that would be designed by Mary would usually be blue.

(51) seems odd for the following reason: The subjunctive marking of the relative clause verb indicates that the eventualities quantified over are located at non-actual worlds, while the indicative marking indicates that they are located at the actual world. In (52), there is no such clash: Both verbs indicate that quantification is over a set of eventualities that are located at non-actual worlds. Further research could include the comparison of the exact conditions for the ungrammaticality of (51) with the interval resolution strategy as presented in the preceding sections.

In Endriss and Hinterwimmer [in preparation] we show that the interval resolution strategy in tandem with the fact that temporal Q-adverbs such as *usually* are only able to quantify over temporally scattered eventualities (in the following referred to as the *coincidence constraint*, cf. Zimmermann [2003] for a related constraint for the interpretation of the adverb *occasionally*, based on Lasersohn [1995]) also accounts for contrasts like the following:

(53)   The people that gave a talk at the conference on kangaroos usually were intelligent.

(54)   *The people that listened to Peter's talk at the conference on kangaroos usually were intelligent.

(55)   The people that listened to Peter's talk at the conference on kangaroos were intelligent for the most part.

In (53), quantification ranges over the parts of the maximal sum eventuality the agent of which is the maximal plural individual of people that gave a talk at the conference on kangaroos. The sentence means that most of those parts are also part of the sum eventuality of being intelligent. (We assume that adverbial quantifiers may not only take sets, but also genuine plural objects as their arguments, cf. the discussion of determiner quantifiers in Matthewson [2001]). A natural partition of the maximal sum eventuality would be the division into eventualities with a different agent each (cf. Nakanishi and Romero [2004]) which in turn accounts for the quantificational variability effect. In (54), however, the maximal sum eventuality introduced by the relative clause verb consists of parts that necessarily coincide temporally, as there is only one talk by Peter at the conference on kangaroos. According to the interval resolution strategy, the running times of the parts of the matrix eventuality quantified over by the Q-adverb have to be set to the respective running times of the parts of the relative clause events. As a result of this, the running times of the eventualities quantified over also coincide, and the coincidence constraint is violated.

As has been pointed out in Nakanishi and Romero [2004], adverbs such as *for the most part* behave differently in this respect. To these adverbs, any plural eventuality whatsoever is welcome (cf. (55)).

In Cohen [2001], Greenberg [1998], Greenberg [2002], and Greenberg [2003], it is shown that there are crucial differences between singular indefinites and bare plurals with respect to generic interpretations. It also seems that bare plurals do not have to obey the interval resolution strategy to the same degree as singular indefinites.

(56)   Cars that were sold in the eighties are usually blue.

For some speakers, (56) seems to be better than (2) which can only be due to the singular/plural contrast of the subject.[14]

Apart from singular/plural contrasts, word order seems to play a role in the interpretation of the respective adverbially quantified sentences[15]:

(57)   Usually, a car that was sold in the eighties is blue.

Sentence (57) is clearly much better than (2).

In future work, we plan a deeper investigation of these phenomena as well as an in depth comparison of the behaviour of singular and plural indefinites with respect to the interval resolution strategy.

Bibliography

E. Bach. The Algebra of Events. *Linguistics & Philosophy*, 9:5–16, 1986.

G. Chierchia. Individual Level Predicates as Inherent Generics. In G. Carlson and F. J. Pelletier, editors, *The Generic Book*, pages 176–223. University of Chicago Press, Chicago, 1995.

A. Cohen. On the Generic Use of Indefinite Singulars. *Journal of Semantics*, 18:183–209, 2001.

H. de Swart. *Adverbs of Quantification: A Generalized Quantifier Approach*. Garland, New York, 1993.

M. Diesing. *Indefinites*. Linguistic Inquiry Monographs. MIT press, Cambridge (Mass.), 1990.

C. Endriss and S. Hinterwimmer. Quantificational Variability Effects with Plural Definites. Manuscript, Universität Potsdam and Humboldt-Universität zu Berlin, in preparation.

---

[14] We would like to thank Jason Stanley who pointed out the contrast to us.
[15] This example is also due to Jason Stanley.

Y. Greenberg. Temporally Restricted Generics. In D. Strolovitch and A. Lawson, editors, *Proceedings of SALT VIII*, Cornell University, 1998. Cornell Linguistics Circle.

Y. Greenberg. Two Kinds of Quantificational Modalized Genericity, and the Interpretation of Bare Plural and Indefinite Singular NPs. In B. Jackson, editor, *Proceedings of SALT XII*, Cornell University, 2002. Cornell Linguistics Circle.

Y. Greenberg. *Manifestations of Genericity*. PhD thesis, Outstanding Dissertations in Linguistics, New York, 2003.

I. Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst, 1982.

E. Herburger. *What counts. Focus and Quantification*. MIT press, Cambridge (Mass.), 2000.

H. Kamp. A Theory of Truth and Semantic Representation. In J. A. G. Gronendijk, T. M. V. Janssen, and M. B. J. Stokhof, editors, *Formal Methods in the Study of Language*, pages 277–322. Amsterdam: Mathematisch Centrum, University of Amsterdam, 1981.

H. Koopman and D. Sportiche. The Position of Subjects. *Lingua*, 85:211–258, 1991.

A. Kratzer. Stage-Level Predicates and Individual-Level Predicates. In G. Carlson and F. J. Pelletier, editors, *The Generic Book*, pages 125–175. University of Chicago Press, Chicago, 1995.

M. Krifka. Focus and the Interpretation of Generic Sentences. In G. Carlson and F. J. Pelletier, editors, *The Generic Book*, pages 238 – 264. University of Chicago Press, Chicago, 1995.

M. Krifka, F. J. Pelletier, G. N. Carlson, A. ter Meulen, G. Chierchia, and G. Link. Genericity: An Introduction. In G. Carlson and F. J. Pelletier, editors, *The Generic Book*, pages 1 – 124. University of Chicago Press, Chicago, 1995.

P. Lasersohn. *Plurality, Conjunction and Events*. Kluwer, Dordrecht, 1995.

A. Lenci and P. M. Bertinetto. Aspect, Adverbs, and Events: Habituality vs. Perfectivity. In F. Pianesi J. Higginbotham and A. C. Varzi, editors, *Speaking of Events*, pages 245–287. Oxford University Press, Oxford/New York, 1999.

D. Lewis. Adverbs of Quantification. In E. L. Keenan, editor, *Formal Semantics of Natural Language*, pages 3–15. Cambridge University Press, Cambridge, 1975.

D. Lewis. *On the Plurality of Worlds*. Blackwell, Oxford, 1986.

S. Löbner. Why German schon and noch are still duals: A reply to van der Auwera. *Linguistics and Philosophy*, 22:45–107, 1999.

L. Martí. *Contextual Variables*. PhD thesis, University of Connecticut, Connecticut, 2003.

L. Matthewson. Quantification and the Nature of Cross-Linguistic Variation. *Natural Language Semantics*, 9:2:145–189, 2001.

R. Musan. Tense, Predicates and Life-Time Effects. *Natural Language Semantics*, 5:271–301, 1997.

K. Nakanishi and M. Romero. Two Constructions with Most and their Semantic Properties. In *Proceedings of NELS 34*, Amherst, 2004. GLSA.

B. Partee. Some Structural Analogies Between Tenses and Pronouns in English. *The Journal of Philosophy*, 70:18:601–609, 1973.

B. Partee. Quantificational Structures and Compositionality. In A. Kratzer E. Bach, E. Jelinek and B. Partee, editors, *Quantification in Natural Languages*, pages 541–602. Kluwer, Dordrecht, 1995.

M. Rooth. Indefinites, Adverbs of Quantification and Focus Semantics. In G. Carlson and F. J. Pelletier, editors, *The Generic Book*, pages 265–299. University of Chicago Press, Chicago, 1995.

S. Rothstein. *Structuring Events. A Study in the Semantics of Aspect*. Blackwell, Malden, 2003.

H. Smessaert and A. ter Meulen. Temporal Reasoning with Aspectual Adverbs. *Linguistics and Philosophy*, 27:209–261, 2004.

J. Stanley. Context and Logical Form. *Linguistics & Philosophy*, 23.4:391–434, 2000.

J. Stanley. Nominal Restriction. In G. Preyer and G. Peter, editors, *Logical Form and Language*, pages 365–388. Oxford University Press, Oxford, 2002.

Z. Vendler. Verbs and Times. *The Philosophcal Review*, LXVI:143–160, 1957.

K. von Fintel. *Restrictions on Quantifier Domains*. PhD thesis, University of Massachusetts, Amherst, 1994.

M. Zimmermann. Pluractionality and Complex Quantifier Formation. *Natural Language Semantics*, 11:3:249–287, 2003.

Cornelia Endriss
Universität Potsdam
Institut für Linguistik
Postfach 601553
14415 Potsdam
Germany
endriss@rz.uni-potsdam.de
http://www.ling.uni-potsdam.de/~endriss

Stefan Hinterwimmer
Humboldt-Universität zu Berlin
Institut für deutsche Sprache und Linguistik
Schützenstr. 21
10117 Berlin
Germany
stefan.hinterwimmer@rz.hu-berlin.de

Discourse Structure and Information Structure: Interfaces and Prosodic Realization

Ekaterina Jasinskaja, Jörg Mayer and David Schlangen

Universität Potsdam

In this paper we review the current state of research on the issue of discourse structure (DS)/information structure (IS) interface. This field has received a lot of attention from discourse semanticists and pragmatists, and has made substantial progress in recent years. In this paper we summarize the relevant studies. In addition, we look at the issue of DS/IS-interaction at a different level—that of phonetics. It is known that both information structure and discourse structure can be realized prosodically, but the issue of phonetic interaction between the prosodic devices they employ has hardly ever been discussed in this context. We think that a proper consideration of this aspect of DS/IS-interaction would enrich our understanding of the phenomenon, and hence we formulate some related research-programmatic positions.

Keywords: discourse structure, information structure, prosody

1 Introduction

In this paper we review the current state of research on the issue of discourse structure (DS)/information structure (IS) interface. Although a recent special issue of the *Journal of Logic, Language, and Information* (Kruijff-Korbayová and Steedman, 2003) has addressed the same topic, the rapid development in the field calls, in our opinion, for another update. Progress has been made both in the study of specific DS/IS interface phenomena (e.g. Büring, 2003; Umbach, 2004; Hendriks, 2004; Zeevat, 2004) and in the development and formalization of the underlying theoretical concepts (e.g. van Rooy, 2003; Ginzburg and Cooper, 2004). These as well as some previous studies will be summarized and discussed in a systematic way.

In addition, we look at the issue of DS/IS-interaction at a different level—that of phonetics. It is known that both information structure and discourse structure can be realized prosodically, but the issue of phonetic interaction between the prosodic devices they employ has hardly ever been discussed in this context. We think that a proper consideration of this aspect of DS/IS-interaction would enrich our understanding of the phenomenon.

Naturally, new findings raise new questions. The ultimate purpose of this paper is to articulate some research areas and formulate hypotheses that should be investigated in order to supply the missing parts of the overall picture.

The paper has three major parts. In Section 2, we present what we take the notions of discourse structure and information structure to mean in isolation, and how each of these structures impacts prosody. In Section 3, we give an overview of the literature that investigates what is traditionally included in the notion of DS/IS-interface. Finally, we present some ideas on the interaction between DS and IS at the level of phonetics in Section 4.

2   General Remarks

Before we take up the issue how discourse structure and information structure interact, it is necessary to say a few words about what we take these two structures to be in isolation, and how they manifest themselves in speech. We will concentrate on their prosodic manifestation.

2.1   What is discourse structure?

Morphology and syntax seek to characterize the well-formedness of words and sentences; similarly, work on discourse structure attempts to describe the coherence of units larger than the sentence, up to the level of entire texts. This enterprise aims to discover and investigate of elementary discourse units, groups

of units that form larger units, as well as the relations between them, which constitute the hierarchical discourse structure.

Hierarchical discourse structure is motivated largely by three kinds of linguistic phenomena that bear on discourse coherence. First, just like syntactic structure constrains the anaphoric relations within a sentence (cf. binding theory), discourse structure affects the accessibility of antecedents for discourse anaphors—anaphoric expressions not captured by binding theory. Second, it is generally accepted that the meaning of a discourse is more than a sum of the meanings of its sentences. In addition, there are various semantic relations (e.g. causal or temporal relations) that hold between the meanings of individual sentences and groups sentences, and which speakers of a language appear to be able to successfully infer from a text, even when they are not overtly signaled e.g. by discourse connectives. Discourse structure provides the blocks to fill the arguments of such semantic relations, whereas in some theories the relations themselves are considered an essential part of the discourse structure, as well. The third kind of linguistic motivation for discourse structure comes from prosody—and we will discuss this in some more detail in Section 2.2. In brief, the crucial observation is that speakers control a number of global prosodic parameters of speech, such as pitch range, speech rate and pause duration, in a systematic way, and that the way they do it intuitively seems to serve the purpose of grouping single utterances into larger chunks.

The current approaches to discourse structure are often classified into two major trends—the *informational* and the *intentional* approaches.[1] Informational approaches attempt to characterize discourse coherence in terms of semantic relations between the information conveyed by successive units. The approaches differ in what role they grant to discourse structure in the overall architecture of language. According to one view, discourse structure is part of the *conceptual structure* (i.e. semantics) of a text. This view is represented by such theories as

---

[1]   See Zaenen et al. (2001) and Asher and Lascarides (2003) for recent surveys.

the *Rhetorical Structure Theory* (RST, Mann and Thompson, 1988), *Discourse Representation Theory* (DRT, Kamp and Reyle, 1993) as well as *Segmented Discourse Representation Theory* (SDRT, Asher, 1993; Asher and Lascarides, 2003), which combines elements of RST and DRT. In particular, RST and SDRT assume large ontologies of semantic discourse relations and the process of computing discourse coherence boils down largely to linking all utterances of a text to one another with these relations in a sensible way. The other group of informational theories treat discourse structure and discourse semantics along the same lines as sentence structure and semantics. That is, discourse structure is allocated at the level of "discourse *syntax*", coherence is treated on a par with sentence-level well-formedness and discourse processing is parsing. The most prominent theories that implement this approach are the *Linguistic Discourse Model* (LDM, Polanyi, 1988) and the *Discourse Tree Adjoining Grammar* (D-TAG, Gardent, 1997; Webber et al., 2003). What unites all the informational approaches, is that they reject using theoretical constructs, such as the speakers' intentions, that reach beyond the text's "syntax" and semantics.

By contrast, the *intentional* approach emphasizes the pervasive role of the speakers' *plans* and *intentions* for discourse coherence. Within this approach, the hierarchy of discourse segments results from hierarchically organized intentions, or *discourse goals* (e.g. Grosz and Sidner, 1986). The idea is that each discourse segment fulfills a goal (conveying certain information, urging the hearer to perform a certain action), and the structural relations between segments reflect the relations between goals. Here, coherence correlates with the efficiency with which a discourse serves the goals of the communication participants. Within the intentional trend, one group of studies will be discussed in particular detail, cf. Section 3.2. These studies represent discourse goals as *Questions under Discussion* (QUD), which the speakers try to answer cooperatively (e.g. Klein and von Stutterheim, 1987; van Kuppevelt, 1995a; Ginzburg, 1996a; Roberts, 1996; Büring, 2003). This framework is particularly interesting

in the current context, since it pays the most attention to information-structural phenomena.

In addition to the major philosophical separation into information and intention orientations, approaches to discourse differ in the richness of the postulated discourse structures, which correlates only partly with the attitude towards intentions. First, the minimum of what a discourse structure is assumed to incorporate is *discourse constituency*, i.e. the information that some discourse segments belong closer together than others, forming larger segments. All the theories mentioned above assume some kind of discourse constituency.

Second, most theories make a further distinction between *subordination* and *coordination* of discourse segments.[2] This distinction affects the accessibility of referents for anaphora in a crucial way (Grosz and Sidner, 1986; Asher and Vieu, 2003). In place of subordination vs. coordination, RST makes a distinction between *mononuclear* and *multinuclear* relations. Although the original motivation for this opposition was rather different (nuclearity is supposed to reflect the communicative weight of an utterance: in a mononuclear relation one sentence is more central and the other one is more peripheral with regard to the message of the text; in a multinuclear relation each part has equal communicative weight), a number of recent studies have shown that the distinction has impact on anaphora resolution (Cristea et al., 1998) and the generation of anaphoric pronouns (Grüning and Kibrik, 2002). In terms of their influence on anaphora, mononuclear relations are roughly comparable to subordination, and multinuclear relations to coordination. The only theories that do not or not systematically make the distinction between subordination and coordination are the ones in the QUD trend, which could be simply due to the fact that the issues of anaphora resolution were not in the center of attention within this frame-

---

[2] In classical DRT, there is strictly speaking no counterpart to coordination: subordination corresponds roughly to the embedding of DRSs, whereas what is viewed as coordinated discourse segments in other theories, would map onto expressions that belong to the same DRS in DRT.

work. However, the few speculations by the proponents of this approach that are found in the literature suggest the general intention to reconstruct the subordination/coordination opposition from the relations between the underlying discourse goals (questions), rather than postulating two kinds of relations between utterances, cf. pp. 184–185.

Finally, some theories assume much richer ontologies of discourse relations than just coordination vs. subordination. These ontologies make further distinctions between relations according to their semantics or pragmatics. E.g. subordinating relations include: *Elaboration*, *Explanation*, *Restatement*, *Comment* etc.; the coordinating relations typically include *Contrast*, *Narration* (or *Sequence*), *Parallel*, *List* etc. For instance if an *Explanation* relation holds between utterances u_1 and u_2 ($Explanation(u_1, u_2)$), the utterance u_2 explains why the event introduced in u_1 happened, e.g. by mentioning the causes of that event. Whereas if u_1 and u_2 are connected by *Contrast*, the speaker intends to compare u_1 and u_2 with respect to their similarities and differences. The approaches that adhere to this rich notion of discourse structure are, for instance, RST, SDRT, LDM and D-TAG. Other proposals in discourse theory reject the idea of encoding such semantic relations as ontological primitives of discourse structure, and either view them as epiphenomenal wrt. speakers' intentions, or acknowledge their existence only to the extent to which they are signaled by overt cues, or are inferable from lexical and world knowledge associated with individual sentences.

We are not going to take sides in this discussion. In this survey we would like to include studies that deal with the interaction of information structure with any of the aspects of discourse structure mentioned above: discourse constituency, coordination vs. subordination, specific discourse relations, as well as hierarchically organized discourse goals. However, it should be noted that we are primarily concerned with one-speaker discourse, i.e. structural relations between utterances in a monologue or within a single dialogue turn. Therefore we have not mentioned turn taking as a further parameter that imposes struc-

ture on discourse. Of course, we will not completely avoid discussing dialogue issues, especially in connection with the question-answer relationship and various question-question relations, which play an important role both in discourse structuring (cf. Section 3.2) and information structural issues. However, we do not pretend to cover the full range of discourse structural, information structural or prosodic phenomena in dialogue. We will touch on these issues only to the extent that they help us understand how DS and IS interact in monologue.

2.2 How is discourse structure realized prosodically?

Recent studies regarding the correlation between discourse structure and prosody differ with respect to the applied discourse model. Three types of approaches can be distinguished: text-oriented approaches, intuitive discourse analysis, and theoretically motivated discourse structures. Studies of the first type (Lehiste, 1975, 1979; Sluijter and Terken, 1993; van Donzel, 1999) equate the structure of written text with discourse structure and concentrate the prosodic analysis on paragraph boundaries in read speech ('paragraph intonation'). Studies of the second type (Ayers, 1994; Venditti and Swerts, 1996; Swerts and Geluykens, 1993; Nakajima and Allen, 1993) are based on discourse or topic models which are intuitively adopted to the specific material of the study. These models consist of very simplified sets of discourse units and relations which are not formally defined or derived from established discourse theories. The third approach, which we consider as the most promising, is to apply one of the independently developed discourse theories and to examine the correlation between the various theoretically identified concepts and the prosodic features of spoken discourse. Most studies of the third type use the intentional approach by Grosz and Sidner (1986), cf. Grosz and Hirschberg (1992); Passonneau and Litman (1993); Hirschberg and Nakatani (1996). But other theories are applied as well in recent studies, namely RST (den Ouden et al., 2002) and SDRT (Mayer, 1999;

Möhler and Mayer, 2001). This third approach allows a far more fine-grained analysis of discourse prosody, increases result comparability, and possibly leads to a generalized model of the interface between discourse structure and prosody.

After these basic remarks regarding the discourse theoretical background of the relevant literature, we will now summarize the results of the above mentioned studies. The two most important prosodic means for structuring longer utterances, which are reported in numerous studies, are pitch range and pause duration. The pitch range is a property of an intonational phrase and defines a subdivision of the total range of fundamental frequency variation of a given speaker. The pitch range can vary in width (e.g. expanded, normal, compressed) and in position relative to the total range (e.g. high, mid, low). It is the reference frame for local tonal events like pitch accents and boundary tones. For instance, a high tone is realized higher in a phrase with expanded pitch range compared to a high tone in a phrase with compressed pitch range. In general, most studies agree that expanded pitch range correlates with the introduction of new discourse topics and sub-topics or with the beginning of a paragraph, respectively. Compressed pitch range, on the other hand, signals the end of a paragraph or the closing of a (sub-) topic. Furthermore, some studies assuming hierarchical discourse structures showed that width and position of the pitch range correlate significantly with the depth of embedding of discourse units (Ayers, 1994; den Ouden et al., 2002; Mayer, 1999; Möhler and Mayer, 2001).

Similar results are reported for the duration of silent pauses. Pauses are longer before units introducing new discourse topics. The shortest pauses appear between intonational phrases dealing with the same topic (Grosz and Hirschberg, 1992; Swerts and Geluykens, 1993; Hirschberg and Nakatani, 1996; Swerts, 1997). As with the pitch range, den Ouden et al. (2002) again showed a strong correlation between pause duration and depth of embedding. However, pitch range (Ayers, 1994) and, even more, pause duration as prosodic correlates of discourse structure depend undoubtedly on speaking style. The clearest results

are obtained from read speech using professional speakers. In a study by Gustafson-Capková and Megyesi (2002), this group yielded results where every silent pause of considerable length was correlated with a topic change, while in the spontaneously speaking group only 34% of the pauses correlated with a topic change.

Apart from pitch range and pause duration, two additional prosodic parameters are considered as relevant for the organization of spoken discourse—speaking rate and intensity variation—but so far, with less convincing results. Concerning speaking rate, Hirschberg and Nakatani (1996) showed that topic-final phrases are produced faster as compared with phrases within the same topic. In contrast, a decrease of speaking rate in phrases preceding a topic change was reported in the study by Smith et al. (2002). A possible explanation for the diverging results could be speaking style again: Hirschberg & Nakatani used spontaneous speech while Smith et al. used read speech. Den Ouden et al. (2002) didn't find any connection between speaking rate and topic structure or depth of embedding, but they found a strong correlation between speaking rate and the nuclearity of discourse segments. Nuclear segments, i.e. segments which are more important concerning the overall coherence of the discourse than others, were realized more slowly. Hirschberg and Nakatani (1996) also reported decreased intensity for topic-final phrases as compared to non-final phrases. These results were confirmed in a study by Herman (2000).

At the end of this section on the prosodic realization of discourse structure we would like to point out that surprisingly little is known about the perceptual relevance of discourse prosodic features. Are these features evaluated by the hearer and if they are evaluated, how are they evaluated? Do we for example perceive pitch range variations gradually or categorically? Is it possible to resolve ambiguous discourse structures with the aid of prosody alone? Are certain prosodic features more important than others? What we know is that synthesized speech with paragraph intonation sounds more natural than without

it (Sluijter and Terken, 1993) and that hearers make use of melody and pauses to identify major discourse units (Swerts and Geluykens, 1994). We assume that not only the research on production aspects of discourse prosody but also the research on perceptual aspects will profit from the integration of formal discourse semantic models and experimental phonetic research as proposed in this paper and in the studies by den Ouden et al. (2002) and Möhler and Mayer (2001).

2.3   What is information structure?

In this paper, we adopt a common view that information structure is constituted by a set of features such as [±F], focus; [±T], topic; [±CF], contrastive focus; [±CT], contrastive topic, etc., defined for all the syntactic constituents of a sentence. The distribution of these features in a syntactic tree affects, on the one hand, the prosodic realization of the sentence (e.g. Selkirk, 1995).[3] On the other hand, these features have their specific semantic/pragmatic interpretations, which affect the presuppositions and sometimes the truth conditions of the sentence, and most crucially, constrain the set of contexts in which the utterance is felicitous (Rooth, 1985, 1992; Schwarzschild, 1999; Büring, 1997). Below, we briefly introduce three most well-known contextual effects of focus ([±F]): felicity of question-answer pairs (Section 2.3.1); felicity and truth conditions of utterances that involve alternatives (Section 2.3.2); and contextual newness vs. givenness (Section 2.3.3). These IS concepts and phenomena will be relevant for our discussion of IS/DS-interface in Section 3.

But before concentrating on these issues, some reservations have to be made. First of all, of course, the approach to information structure we adopt here is only one of a great many proposed in the literature. However, a comprehensive review goes beyond the scope of this paper. The reader is referred to van Kup-

---

[3] How exactly information structure is realized prosodically, will be recapitulated in more detail in the next section (2.4).

pevelt (1999) and Kruijff-Korbayová and Steedman (2003) for recent surveys and further references.

Second, we restrict our discussion to *sentential* information structure, i.e. focus and topic features defined and interpreted not higher than at the sentence level. Further, we will remain open as to whether there is a single information-structural partition of a sentence, or whether foci and topics can be embedded in other foci and topics. It should be noted though that the issue is not at all trivial and has been discussed in the literature, cf. Partee (1996); Komogata (2003). In the current context it is particularly interesting to note that, at least in some languages, we find cases that suggest that foci can be interpreted above the sentence level, i.e. one full sentence (a discourse) may serve as focus with respect to another sentence, which is then completely deaccented, cf. (Kodzasov, 1996). This opens a completely different perspective on possible DS/IS interactions—topic and focus features taking scope over discourse constituents. However, we will not go further into this issue in this paper.

Finally, we are primarily interested in information-structural categories motivated by prosodic phenomena. It is well known that also syntax, morphology and certain lexical items (particles) are sensitive to information structure, and that frequently, the IS categories motivated by morphology and syntax and IS categories motivated by prosody are not the same (Vallduví and Vilkuna, 1998). Therefore we resrict our discussion to the latter.

2.3.1   Focus in question-answer pairs

One of the contexts most universally acknowledged as a diagnostic for focus are question-answer pairs. The constituent of the answer that corresponds to the *wh*-word of the question should be focused (F-marked), and hence bear a nuclear accent (indicated by small caps in the examples below). This makes (1-a)–(2-a) and (1-b)–(2-b) felicitous question-answer pairs, whereas (1-a)–(2-b) and (1-b)–

(2-a) are infelicitous.

(1)   a.   Who did Mary vote for?
       b.   Who voted for John?

(2)   a.   Mary voted for [JOHN]$_F$.
       b.   [MARY]$_F$ voted for John.

This simple example illustrates the contextual function of focus in answers to questions. Some linguists propose to reduce the pragmatics of focus to this function alone. They postulate that every sentence answers a question, but the question may be implicit, whereas further constraints at the level of discourse structure regulate which implicit questions are admissible in which contexts. Basically, this approach postpones the analysis of most focus-related contextual phenomena to the level of discourse structure. This will be discussed in some detail in Section 3.2. By contrast, other approaches attempt to capture a broader range of the contextual effects of focus at the level of IS by giving more general definitions of focus pragmatics, cf. below.

2.3.2   Focus and alternatives

ther well-known approach (cf. Rooth, 1985, 1992) sees the pragmatic function of focus in highlighting information that is contrary to some *alternative(s)*, anaphorically recoverable from the context. That's why B is a felicitous denial of A in (3), whereas B' is not: *voted for Bill* is a legitimate focus alternative for the VP *voted for [JOHN]$_F$*, but not for *[VOTED]$_F$ for John*.

(3)   A: Mary voted for Bill.
       B: No, she voted for [JOHN]$_F$.
       B': # No, she [VOTED]$_F$ for John.

The interpretation of focus in terms of alternatives proved particularly fruitful as a way to account for the semantics of sentences with so-called focus-sensitive particles, such as *only*. Roughly speaking, *only* conveys the idea that the focus-alternatives of the expression in its scope do not hold. For instance, (4-a) implies that Mary did not vote for Bill, or George; whereas (4-b) says that Mary did not do anything else for John, e.g. she did not campaign for him, but she could have voted for other people, as well.

(4)   a.   Mary only voted for [JOHN]$_F$.
       b.   Mary only [VOTED]$_F$ for John.

The alternative-based approach also gives an adequate account of the effect of focus in question-answer pairs, as well as a whole range of other contexts (cf. Rooth, 1992).

2.3.3   Givenness

Finally, we would like to mention the proposal by Roger Schwarzschild (1999), according to which, a specific pragmatic function is carried by the *un*focused material, i.e. the [−F]-constituents, rather than [+F], as in the approaches discussed above. This is motivated by the observation that [−F]-constituents can be uniformly interpreted as *given*, or 'anaphorically recoverable', whereas [+F] needs, according to Schwarzschild, at least three distinct pragmatic definitions. Two of them were presented in the previous sections: it is (a) focused material as 'replacing the *wh*-element in a presupposed question' (cf. 2.3.1); and (b) focused material as 'being contrary to some predicted or stated alternative' (cf. 2.3.2). The third definition of [+F] relates it to 'textually and situationally non-derivable information', such as *make* in B's answer in (5).

(5)   From (Schwarzschild, 1999, p. 142):

A: Why don't you have some French TOAST?

B: I've forgotten to MAKE French toast.

According to Schwarzschild (1999) this multitude of (unrelated) definitions is due to a redundancy in the conceptualization: [−F] (*given*) and [+F] (*new*) really are complements, and so only one notion is needed in the theory. His solution is thus to make the notion of *givenness* elementary, and to link only this to intonation.

We will gloss over the technical details here,[4] and focus on the main points: (6) shows the (simplified) definition of *given* in this approach; (7) links it to F-marking; and (8) gives the additional constraint that is needed to restrict F-marking (since (6) on its own does not say anything about the status of focused information).

(6) *Given* (simplified)

An utterance U counts as *given* iff it has a salient antecedent A which entails the non-F-marked parts of U.

(7) *Givenness*

If a constituent is not F-marked, it must be *given*.

(8) *AvoidF*

F-mark as little as possible, without violating *Givenness*.

Augmented with constraints that relate F-marking to accentuation, the theory makes plausible predictions wrt. accent placement, i.e. which constituents of a sentence must, may, or must not be accented in which contexts.

---

[4] Just briefly: to be able to define *given* as 'being entailed by salient parts of the previous discourse', Schwarzschild (1999) defines a semantic operation called 'existential type shifting' that takes arbitrary parts of (the meaning of) antecedent utterances to the type of formulae which can then be tested for whether they entail non focused parts (similarly type-shifted to type t) of (the meaning of) the new utterance.

Thus we have briefly sketched three closely related approaches to the pragmatic interpretation of focus features [±F]. In spite of some differences in the details, all three belong to the same tradition and reflect the notion of information structure presupposed in this paper. The pragmatic definitions of other IS-features, such as topic, discussed in sections below, build up on the presented view of focus.

2.4   How is information structure realized prosodically?

In the previous section we discussed the pragmatic interpretation of focus features, now we turn to their prosodic realization. The best-studied prosodic means of marking focus is the placement of the (nuclear) accent. Some word of the focused constituent receives an accent (cf. the words in small caps in the examples (2)-(5) in the previous sections). This word is then called the *focus exponent*. Roughly, if the focused constituent coincides with the focus exponent, one talks about *narrow focus*; a focused constituent that is larger than its focus exponent is called *broad focus*. In the latter case, special syntactic and pragmatic rules regulate which word of the focused constituent is accented (Selkirk, 1995; Schwarzschild, 1999).

In many languages, e.g. English and German, focus exponents receive pitch accents. However, it is also very common that besides accent position, focus constrains accent type or tune: focus-related pitch accents on new material are usually falling accents (HL-sequences in analyses based on the tone sequence model by Pierrehumbert, 1980), whereas topic accents are often—for example in German—rising accents (LH-sequences, cf. Büring, 1997). Furthermore, languages can prosodically differentiate between broad and narrow focus by means of categorically distinct pitch accent realizations (e.g. Frota (2000) for Portuguese). While it is a standard view that prosodic prominence on phrase level is expressed essentially by local pitch movements, we know little about

the supporting function of other prosodic parameters like e.g. the highlighting of focused material by filled or silent pauses (Arnold et al., 2003; Horne et al., 2004) or the variation of speaking rate to differentiate between given and new material.

Another interesting area in the field of prosody and information structure is the mechanisms of prominence reduction. In Germanic intonation languages like Dutch contextually given material is deaccentuated (complete deletion of pitch accents), while in Romance languages like Italian pitch accents on given material are realized, but with significantly reduced accents range compared to new material (Swerts et al., 1999). In Swedish, a language with lexical accent, pitch accents on contextually given material can also be realized if they occur early in the sentence, but they differ from accents on new material in peak-alignment (Horne et al., 1999). Focus exponents in the scope of a focus-sensitive particle are deaccentuated (deletion of pitch accent), if they occurred already in the identical construction in the immediately preceeding context ('second occurrence focus'), but are still marked prosodically prominent by means of durational and intensity features (Beaver et al., 2002).

The last area we would like to sketch out briefly is prosodic phrasing. While in neutral constructions phonological phrasing is a mere reflection of syntactic structure, both levels can vary independently under the influence of information structure. For example, word order variations due to information structural constraints can result in identical phonological phrasing patterns. On the other hand, identical syntactic structures can be phonologically implemented with different numbers of phrases to meet information structural requirements, such as, for example, narrow emphatic focus.

2.5   Interim conclusions and further questions

In the previous sections we formulated our assumptions about the essence of discourse structure and information structure and reviewed recent findings about their prosodic realization. In our opinion, the concepts and facts presented above suggest (at least) two possible ways in which DS and IS could interact.

First, DS and IS could interact at the level of pragmatics. On the one hand, the pragmatics of topic and focus contributes to the appropriateness of an utterance in a given context. On the other hand, discourse structure provides a highly structured representation of the context. Depending on the discourse structure, not all parts of the context might be equally relevant for the distribution of topic and focus features in a particular utterance. In turn, information structure could impose constraints on how the utterance is connected to the context. What these constraints are like, whether these constraints are "direct" or mediated by other structures, what aspects of discourse structure interact with which information-structural features—all these questions and the relevant discussion in the literature will be recapitulated in Section 3.

Second, it is obvious that information structure and discourse structure (see Section 2.2) use the same prosodic devices, namely pitch, durational and intensity parameters, to express different kinds of structures, relations, and prominence. This proximity suggests phenomena of interaction and conflicting requirements. We formulate some related hypotheses in Section 4.

3   DS and IS Interact at the Level of Pragmatics

In this section, we review the current state of research on the issue how discourse structure and information structure communicate at the level of (semantic/pragmatic) interpretation. The field has made substantial progress in recent years. Two major theoretical trends seem to be emerging. In one of them, the

interaction between discourse and information structure is assumed to be mediated by the speakers' joint intention to discuss a certain issue, or a *Question under Discussion* (QUD). The other trend does not make use of this construct and tries to relate DS and IS more or less directly. Since this approach is in some sense simpler, we will start by presenting it in Section 3.1. The approach based on QUD will be discussed in Section 3.2. The QUD framework is conceived in such a way that practically any phenomenon related to information structure (e.g. those discussed in Section 2.3) has to be analyzed as a DS/IS interface phenomenon. Therefore we will dwell somewhat longer on the general architecture of that approach.

3.1 Relational approaches

In this subsection we discuss the (at first glance seemingly unconnected) approaches by Schwarzschild (1999), Asher and colleagues (in the framework of SDRT, see in particular Asher and Txurruka, 1995) and Nakatani (1997). What unites these approaches for our purposes is the common assumption that there are coherence relations that directly link utterances (without recourse to structures like QUDs) and that have an influence in licensing information structure—be they the relations of centering theory that hold between forward- and backward-looking centers (as in Nakatani, 1997), rhetorical relations (as in SDRT), or the relations between anaphor and antecedent, mediated by rhetorical relations (as in Schwarzschild, 1999). We begin with a discussion of the approaches and then look at some phenomena that have been dealt with in these approaches.

3.1.1   The approaches

Givenness as anaphora: Schwarzschild (1999)

In Section 2.3.3 we already introduced the main positions of Schwarzschild's (1999) work, in particular, the notion of *givenness* and its relation to focus marking and accent placement. Here, we address the predictions this approach makes wrt. interactions between accent placement and discourse structure. The part of the proposal that is of the most interest here is the claim that givenness is a form of anaphora, which has to search for a *salient* antecedent, cf. definition (6) in Section 2.3.3. Schwarzschild (1999, p. 165) conjectures that salience is mediated by rhetorical relations (and hence by discourse structure). The author does not mention any particular theory of discourse structure or accessibility, but it would be interesting to test whether this *given*-anaphora behaves in the same way as pronominal anaphora. To give a brief illustration of this question, consider (9) below. The SDRT-conditions on accessibility of antecedents (Asher and Lascarides, 2003) explain why the continuation (e) in this mini-discourse is odd (namely because utterance (e) is in a narrative sequence with (b), which renders the evaluation (c)–(d) of (b) 'inaccessible').

(9)   (a)   Sandy had a great evening.

        (b)   First she called Peter.

        (c)   They talked about her mother.

        (d)   And about some common friends.

        (e)   ??Then she phoned her.

The better acceptability of the intonational contours in (10) below suggests that *givenness*-anaphora underlies weaker constraints than SDRT-accessibility (since (e$''$) is only licensed by attaching to (c), which in SDRT is not available); this however remains to be tested in a more rigorous empirical manner.

(10) (e′) Then she phoned her MOTHER.

(e″) Then she PHONED her mother.

Information packaging in SDRT

Asher and Txurruka (1995) integrates a theory of *information packaging* into the discourse theory SDRT (Asher, 1993; Asher and Lascarides, 2003)—'information *packaging*' (rather than 'information *structure*') being the term that Chafe (1976) introduced (and Vallduví (1992) took up and further developed) to emphasize the function (possibly one among many) that intonation has of indicating to the hearer how the parts of an utterance fit into the context. Asher and Txurruka (1995) make the following assumptions: (a) both the topic/comment-dichotomy and the focus/background-dichotomy have a role in a theory of information packaging; (b) one of the dichotomies becomes *dominant* in a given utterance (in a given context), resulting in utterances being either *focus-dominant* or *topic-dominant*; (c) information packaging is a pragmatic notion, which is only partially determined by intonation or syntax; (d) *rhetorical relations* connecting utterances can constrain the informational structure of their relata (or conversely, a given informational structure can be used as evidence for inferring that a certain rhetorical relation holds).

This already represents the essential ideas of the approach; in the following, we will try to illustrate them a bit more. The distinction in (b) above is motivated by the difference between questions as in (11-a), which set up a *topic* for discussion, and those in (11-b), which set up a *situation* for discussion—and in this approach, this is what the discourse context has to provide, namely either an entity (a topic) or a situation (a background).

(11) a. A: What about John? What did he do?

B: JOHN (↗) talked to MARY (↘)

b.   A: Who did John talk to?

B: John talked to MARY.

According to the authors, certain rhetorical relations are sensitive to kinds of information packaging; e.g. relations connecting the utterances to the first type of question in (11) above require topic-dominant structures, as probably do relations like *Narration* and *Elaboration* (although this is not developed in the paper), whereas, for example, *Corrections* require a focus-dominant structure, just like the second kind of question-answer-pair above.

Now, what are those information packaging-sensitive constraints? While the technical details of the proposal are involved (making use of partial isomorphisms between semantic representations, i.e. the SDRSs), the main idea is simple: the focused element in the new information must be mapped to an element in the antecedent utterance, and the remaining parts of both utterances (the results of abstracting over these elements) must be logically equivalent (possibly in a non-monotonic logic that allows for 'normality'-modalities).[5,6] An example will hopefully make this clearer:

(12)   A: Who came to the party?

B: JOHN came to the vernissage.

If we map "John" to "who", we get "came to the party" as the remaining elements of the first sentence, which is equivalent to "came to the festivity" (and

---

[5]   This amounts to essentially the same idea as in Schwarzschild's definition of *given*, only in a more direct way: where Schwarzschild ensures through use of type-shifting and existential closure that the antecedent and the given-elements in the utterance are of the right type for the first to entail the second, in this approach, the antecedent utterance itself (or rather, its logical form, LF) and a constructed utterance (LF) as described above are related by the entailment relation.

[6]   In Asher and Txurruka (1995) only focus-dominant-sensitive rhetorical relations are worked out in detail; in later work in SDRT (e.g. Asher and Lascarides (2003)), the distinction between focus-dominance and topic-dominance is not even mentioned.

this shows why logical equivalence is needed, namely because the situation can be referred to differently in either sentence). For (this type of) question, the authors can then additionally require that the focus be mapped onto a *wh*-element; in corrections such as (13) below, the mapping is between elements of the same type (the condition on the background is the same).

(13)   A: John came to the party.
         B: No, it was PETER who came to the fete.

Note that contrary to Schwarzschild's motivating assumptions, this approach does not see it as problematic that there is no uniform notion of *newness* (as the relation that connects focused material to an antecedent): there are different rhetorical relations that hold between the utterances in any case, and it is only natural for them to impose different constraints on their relata. Information-packaging constraints on the relation *Narration* aren't worked out in this approach as far as we can see, so there is no developed treatment of our accessibility example above (9), but it should be clear that the mechanism is there in this approach to explain the observed phenomena.

Intonation and Centering Theory

The last approach we will discuss here, that of Nakatani (1997), sticks out a bit in that it is not an attempt to formalize constraints on accent placement, but rather is an analysis of empirical data, in the framework of a specific theory of discourse structure. More precisely, she looked at referring expressions, and correlated features of their form (pronouns vs. full forms), the grammatical function they play in the utterance (subject vs. object), and intonation (prominent vs. non-prominent) to their function according to *Centering Theory* (Grosz et al., 1995). As we will see, this analysis can be seen as keeping the link between intonational (non-)prominence and newness (givenness), while giving a

more nuanced analysis of the second set of terms.

The purpose of Nakatani's study was to test "the general claim about the accentuation of given/new information" which "predicts that (1) pronouns are unaccented and full noun phrases and proper names are accented, and (2) subjects are unaccented and direct objects are accented" (Nakatani, 1997, p. 140). While Nakatani found a supporting tendency, there were also many exceptions to this rule, which she claims can be explained by looking at the status of these expressions with respect to centering theory. Before we can discuss this, we will briefly review the fundamentals of this theory.

Centering looks at patterns of referential connections between utterances, i.e. at how subsequent sentences in a (coherent) discourse keep referring to the same entities or introduce new ones, and formulates rules about which kinds of referential configurations produce "better", i.e. more coherent, discourse. To give an example, centering predicts, corresponding to intuitions, that the discourse in (14-a) is relatively more coherent than that in (14-b), which both mention exactly the same entities, only in different order.

(14)   a.   Peter went to a music store.
                He really liked that store.
                He bought a piano.

         b.   Peter went to a music store.
                It was a store that he really liked.
                A piano was what he bought there.

To formulate the transition rules, centering looks at two aspects of utterances, namely how they set up entities for further reference in subsequent discourse (the *forward-looking* potential), and how they take up entities from previous utterances (the *backward-looking* aspect). The former aspect is captured by assigning each utterance a list in which the entities that are mentioned in the utter-

ance are ranked according to criteria like the grammatical function they play in the sentence (with subjects being higher ranked than objects, for example), or linear order; this is the list of *forward-looking centers* or C_f. Broadly speaking, this ranking is meant to reflect the likelihood that an entity will be mentioned again in the next utterance, or the preference that it be rementioned; hence the highest ranked member is also called the *preferred center* or C_p. The other aspect is represented in the *backward-looking* center C_b, which is the highest ranked member of the previous utterance's C_f that is realized in the current utterance. Differences in perceived coherence (as in example (14) above) are then explained by ranking movements of configurations of forward and backward-looking centers: e.g. *retaining* a center (as in all utterances in (14-a), where "Peter" remains the highest ranked element of C_f) is preferred over *rough shifting* centers from one utterance to the next (as in (14-b), where the highest ranked element is "Peter", "the store", and "a piano", respectively).

This theory of centering is embedded in a more general theory of discourse structure (Grosz and Sidner, 1986); it suffices here to say that in this theory new discourse segments introduce new *centering spaces* which are put on a stack (i.e. which are removed once a segment is closed, to return to the previous centering configuration). These two elements form the theoretical framework in which Nakatani formulates her analysis,[7] which is presented in Table 1.

Perhaps the only surprising result is the occurrence of prominent pronouns in subject position; according to the author, these were cases like the following (our example):

(15)   Peter likes Sandy.

---

[7] A word on the scope of the analysis. Nakatani (1997) looked at spoken narrative (i.e. not dialogue), and centering has indeed been developed for monologues (but see Taboada (2002) for a review of attempts to extend centering to dialogue), hence there are no immediate predictions of this approach with respect to the examples we have seen so far in the previous sections.

p	gf	f	discourse function
+	S	p	shift local attention to new C_b
+	S	ef	Introduce new global referent as C_p
+	DO	ef	Introduce new global referent
−	S	p	Maintain C_b referent in primary local focus
−	DO	p	Maintain non-C_b referent in primary local focus
−	DO	ef	Maintain referent in global focus

Table 1: *pr* stands for "prominence"; *gf* for "grammatical function"—*S* being subject and *DO* being direct object; and *f* for "form", with *p* for "pronoun" and *ef* for "explicit form".

SHE hates him, though.

However, an important point is made explicit in this analysis (whereas it is implicit in both approaches that have been discussed above), namely that *givenness* or *newness* of a given entity is relative to previous discourse structure, and is not absolute over the discourse as a whole, or in other words, that *salience* has a role to play.

As a point of criticism, it has to be pointed out that it's not entirely clear what exactly Nakatani is talking about: is it nuclear accent, or sentential focus? As far as we can see, she doesn't say whether one element being prominent implies that all other elements are non-prominent, or whether there were configurations where both subject and object were intonationally prominent.

While Nakatani phrases her analysis rather procedurally in terms of instructions to the listener (e.g. "shift attention to"), it should be possible to reformulate her approach so that it is more in line with the ones discussed above (e.g. "intonational prominence on a subject pronoun is licensed if it is a new C_b"), and hence, this difference does not favor or disfavor any approach.

3.1.2 The phenomena

In this section we very briefly mention some discourse configurations where IS has an effect on felicity conditions, and say whether, and if so, how the discussed approaches handle them.[8]

Questions and answers

The well-known IS constraints on questions and answers (already mentioned above), illustrated by (16), are handled by Schwarzschild (1999) through the *givenness* mechanism which handles all discourse effects: in B′ below "Sandy" is not given, and hence its non-prominence violates that constraint.

(16)   A: Who came to the party?
        B: SANDY came to the party.
        B′: *Sandy came to the PARTY.

Asher and Txurruka (1995), on the other hand, have specific constraints for question-answer-pairs, as discussed above.

Contrast

Example (17) illustrates the IS-constraints on contrastive sequences. Again the fact that "speaks" is non-prominent can be explained by Schwarzschild (1999) with his *givenness* constraint. This approach, however, has nothing to say about why the other elements have to be accented with this particular contour. Moreover, to explain the realization of the first sentence, the approach would have

---

[8] Note that all those dialogue acts (answering, correction, clarification) can also be performed with fragmental utterances (e.g. "A: Peter came to the party.—B: No, SANDY.") where the fragment consists only of the focus. This has been noted before (Ginzburg (1999b); Schlangen (2003)), but has not yet been systematically studied, to our knowledge.

to allow *given* to be *cataphoric*, i.e. to take its "ante"cedent from subsequent discourse.

(17)   JOHN (↗) speaks FRENCH (↘).
        BILL (↗) speaks GERMAN (↘).

IS-constraints on *Contrast* are mentioned in passing in Asher and Lascarides (2003), and again the general idea is that this relation requires this particular contour, and conversely that this contour can be used to infer this relation.

Correction and focus

Corrections have already been discussed above; the treatment in the approaches follows the same general lines as for other phenomena.

Clarification and focus

The influence of focus-marking on the interpretation of clarification requests (i.e. questions addressing the question of understanding a previous utterance, including acoustic understanding; cf. Ginzburg (1999a); Schlangen (2004)) like those in (18) has so far not been studied systematically, but it seems that there are constraints on the information structure which work very similarly to those on corrections, and which help determine which elements are being clarified.

(18)   A: Peter hit Sandy.
        B: Peter HIT Sandy? (= "Did you say that what Peter did to Sandy was to hit her?")
        B′: Peter hit SANDY? (= "Did you say that the person Peter hit was Sandy?")

$$Q$$
$$Q_{1-2} \quad Q_{3-4}$$
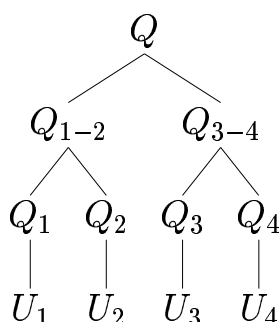$$Q_1 \quad Q_2 \quad Q_3 \quad Q_4$$
$$U_1 \quad U_2 \quad U_3 \quad U_4$$

Figure 1: Abstract QUD-structure of a sequence of utterances $\langle U_1, U_2, U_3, U_4 \rangle$. The utterances are immediately dominated by (i.e. answer) the questions $Q_1, ..., Q_4$, respectively. The dominance relations between questions correspond to subquestion relations. The whole discourse is dominated by the root node Q, i.e. the discourse answers the question Q.

3.2 Questions under discussion

3.2.1 Theoretical background

The so-called *Question under Discussion* (QUD) approach is based on the assumption that *each* utterance in discourse answers an explicit or implicit question (under discussion). If the question is actually uttered by one of the conversation participants, it is *explicit*, and *implicit* otherwise. The relations between utterances are construed as relations between the underlying questions.[9] This contrasts with the assumptions made in other frameworks, such as RST, SDRT, D-TAG, in which discourse relations connect the utterances directly (cf. Section 2.1). We are not going to discuss the specific gain of such a complication in the theory. For the time being it is enough to say that QUD is at present the most prominent framework addressing the relation between discourse structure and information structure as one of its central issues.

It is necessary to note that the QUD approach does not form a coherent framework comparable to SDRT or D-TAG. Rather, it is a collection of partly

---

[9]   A schematic discourse structure (Büring's style, cf. Büring, 2003) is given in Figure 1.

isolated independent proposals, which include Klein and von Stutterheim (1987, 1992, 2001); von Stutterheim and Klein (1989); van Kuppevelt (1995a,b); Roberts (1996, 1998); Ginzburg (1996a,b); Larsson (2002); Cooper (2003); Ginzburg and Cooper (2004); Büring (2003). In this section, we try to formulate something that could be viewed as the "core" of the QUD framework, paying tribute to the individual proposals where possible.

The cue notions of the QUD framework are those of *question*, *answer*, *inquiry strategy* and various relations between questions within a strategy.

Questions and answers:

According to the standard view, a question is identified with the set of its possible answers. This idea was first introduced by Hamblin (1973) in connection with the semantic analysis of interrogative sentences. It was further developed by Karttunen (1977), Higginbotham and May (1981), Groenendijk and Stokhof (1984), among others, covering more and more aspects of linguistic behavior of interrogatives. However, all these proposals were aimed at providing a purely context-independent compositional semantics for interrogative sentences. It was acknowledged that our understanding of a question, and hence our ability to give an appropriate answer is strongly dependent on what is *relevant* in the current situation, but only with the work of Ginzburg (1995) was the idea explicitly integrated into the formal analyses of interrogative semantics. Most of the work in the QUD-framework adopted a notion of quetion based on Ginzburg's context-sensitive approach to the semantics of interrogatives.[10] The idea, or at least one of its aspects, is illustrated in examples (19) and (20), which show two possible answer-set interpretations of the question *Who attended the meeting?*

---

[10] In the work by Klein and von Stutterheim (1987, 1992, 2001) and van Kuppevelt (1995a,b) sets of propositions associated with the nodes of a discourse tree are called *topics*, rather than *questions*. However, all QUD-based proposals seem to agree on the type of object associated with discourse nodes. Therefore, we ignore these terminological differences and uniformly call these objects *questions*.

(19)   a.   John attended the meeting.

        b.   Mary attended the meeting.

        c.   John and Mary attended the meeting., etc.

(20)   a.   Eight linguists attended the meeting.

        b.   Fifteen psychologists attended the meeting., etc.

Which interpretation is actually chosen, i.e. what is currently relevant, seems
to be regulated by two major classes of constraints. First of all, as was argued
by Ginzburg, questions (as discourse goals) depend on the domain-level goals
of the interlocutors. See van Rooy (2003) for a formal account of this relation-
ship in terms of statistical decision theory. Second, questions addressed in a
discourse depend on each other. For instance, the set of propositions associated
with question Q_2 in Figure 1 depends on the sets Q_1 and Q_{1-2}—the nodes Q_2
is structurally related to. This class of constraints is discussed in more detail in
the next section.

Relations between questions:

As mentioned above, in the QUD framework relations between utterances in
discourse are cast in terms of relations between the (implicit) questions they an-
swer. There are two major ways in which such relations are characterized, which
we could dub as *static*, or declarative, and *dynamic*, or procedural. However, it
appears that both approaches can be successfully used to model interdependen-
cies between questions in discourse.

The static approach: Within the static approach, certain relations between
questions play the role of constraints on allowed discourse structures, the most
important being the relation of *subquestion* and the underlying notion of *in-
quiry strategy*. Informally, the set S of questions $\{Q_1, ..., Q_n\}$ is a *strategy* for
answering question Q_0 iff answering all of the questions in S gives an answer

to Q_0. A sequence of utterances $U_1, ..., U_n$ forms a discourse constituent dominated by question Q_0 only if $U_1, ..., U_n$ are dominated by (i.e. answer) the questions $Q_1, ..., Q_n$ respectively which in turn form a strategy for answering Q_0. Various versions of the notion of *strategy* are present in both Klein and von Stutterheim's and van Kuppevelt's work, although they do not use the term.[11] Roberts (1996) seems to have introduced the term in the discourse-structural context. She relates the notion to Groenendijk and Stokhof's question entailment, whereas Büring (2003) has explicitly used it to formulate constraints on discourse structures in a consistently declarative way.

Given the notion of strategy, a *subquestion* is defined as follows: Question Q is a *subquestion* of question Q' iff it belongs to some strategy of answering Q' (cf. Roberts' (1996) and Büring's (2003) notion of relevance of a question). It should be noted though that frequently subquestionhood is given an "absolute", strategy-independent definition: Q is a *subquestion* of Q' iff in order to answer Q', Q has to be answered first (cf. van Kuppevelt (1995a) and Ginzburg's (1996a) notion of *dependent* qestions). However, such a definition appears too strong, because often there are multiple alternative strategies of answering a question that are equally good. For instance, a question like *Who ate what?* can be addressed via at least two strategies: going "by people", cf. Figure 2; or going "by food", cf. Figure 3. Of course, we would like to treat a question like *What did John eat?* as a legitimate subquestion of *Who ate what?*, but strictly speaking, it need not be addressed in order to find out who ate what. Namely, one could do without asking *What did John eat?* if one chose the by-food strategy.

To summarize, within the static approach questions under discussion are used to define the hierarchical discourse structure based on the relation of *sub-*

---

[11] Jan van Kuppevelt's analogue to strategies is the relation of *subquestion* (cf. below) and, as he calls it, the "conjunctive property of subquestions" (van Kuppevelt, 1995a, p. 125, 24ff). Klein and von Stutterheim use the term *referentielle Bewegung* 'referential movement'. They do not give a general definition, but formulate a number of text type specific constraints on referential movement, which can be seen as special cases of strategies in the above sense.
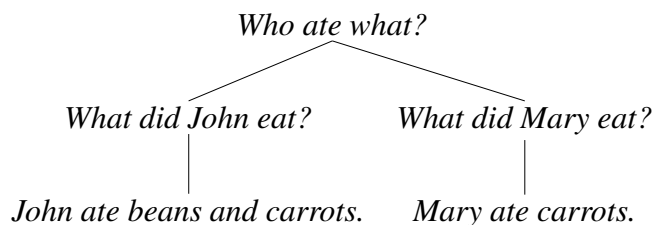
Who ate what?

What did John eat?   *What did Mary eat?*

John ate beans and carrots.   *Mary ate carrots.*

Figure 2: Addressing the question *Who ate what?*: The "by-person" strategy.

Who ate what?

Who ate the carrots?   *Who ate the beans?*

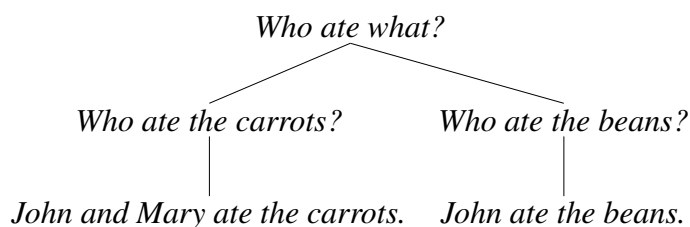*John and Mary ate the carrots.*   *John ate the beans.*

Figure 3: Addressing the question *Who ate what?*: The "by-food" strategy.

question and *strategy*. The possible combinations of question nodes in the resulting tree structure is constrained by the definition of *subquestion* relation, based on Groenendijk and Stokhof's notion of question entailment or Ginzburg's dependent questions. However, as we will see in Section 3.2.2, these constraints are not enough, for instance, to predict the correct distribution of topic accents.[12]

The dynamic approach: A different style of formulating the same constraints is represented by a considerable bulk of work on dialogue (Ginzburg, 1996a,b; Larsson, 2002). Rather than formulating constraints on possible discourse trees, the context of an utterance is modeled by a highly structured *information state* and a set of rules for dealing with it that models the behavior of discourse par-

---

[12] See our discussion of Büring (2003) on pp. 187–189.

ticipants during a conversation. The information state includes, among other things, a stack of questions under discussion (the QUD stack).[13] The question at the top of that stack corresponds to the current question under discussion, whereas the other questions below it correspond to its superquestions.[14] The QUD stack is *updated* and *downdated* by the interlocutors in the course of conversation. If a question is explicitly asked by one of them, it is pushed on the QUD stack, i.e. the QUD stack is updated. A question can also be pushed on the stack, if it is "presupposed" by some utterance, in which case the question has to be *accommodated*.[15] Accommodated questions roughly correspond to implicit questions in the static view. Once a question is on the QUD stack, the conversation participants are committed to address it until it is either answered or determined to be practically unanswerable (cf. Roberts, 1996). However, once a sufficient answer is provided, the question must be popped, i.e. the QUD stack is downdated (cf. van Kuppevelt's (1995a) Dynamic Principle of Topic Termination). When a question is popped off the QUD stack, its immediate superquestion becomes topmost and the procedure is repeated until there are no more questions left (cf. van Kuppevelt's (1995a) Principle of Recency). The QUD stack management in the dynamic approach can be viewed as a way of processing hierarchical discourse trees defined in the static approach.[16]

It should also be noted that in addition to hierarchical relationships between questions and their processing, the QUD-based dynamic analyses of dialogue have concentrated in particular on modeling the distinction between the public and the interlocutors' private part of the information state, and the process of *grounding*. The idea is that an utterance is not *automatically* accepted by the

---

[13] The notion that gave its name to the whole trend in discourse theory.

[14] A superquestion is the reverse of subquestion: if Q_1 is a subquestion of Q_2, then Q_2 is a superquestion of Q_1.

[15] Roberts (1996) argues that it is a general function of focus to introduce a question presupposition.

[16] Note that processing a context-free grammar requires a stack.

interlocutors, i.e. a question is not immediately added to the QUD stack and an assertion does not immediately update the common ground (the set of shared facts). First it has to be understood by all the interlocutors and acknowledged to be relevant to the current question under discussion. Most importantly, the question whether an utterance should be accepted is by itself a discussable issue, and should be modeled in such a way that it can be integrated in the overall hierarchy (stack) of questions. See in particular Ginzburg (1997), Ginzburg and Cooper (2004) and Ginzburg (forthcoming). These issues have not received much attention in the analysis of monologue, whereas, as we already noted, the notion of grounding in a monologue is probably key for a QUD-based definition of subordination.[17]

Relations between declarative utterances

Relations between declarative utterances in a monologue have received much less attention in the QUD framework than in "relational" theories of discourse, such as RST or SDRT. Indeed, it seems that what the hierarchy of questions gives us is discourse constituency, i.e. which utterances belong closer together than others (one strategy vs. different strategies), but it does not seem to provide any distinction between coordination and subordination, or various semantic relations such as *Elaboration*, *Explanation*, and *Contrast*.

At present, there is no systematic account of these aspects of discourse structure in the QUD framework, but there are a few ideas circulating on how these issues could be approached. A promising starting point is van Kuppevelt's (1995a; 1995b; 1996) classification of unsatisfactory answers and the notion of *grounding* in the dynamic trend (Ginzburg and Cooper, 2004). A speaker, even a speaker of a monologue that gets no on-line feedback from the audience, sometimes has reasons to assume that the utterance he has produced gives an

---

[17] See Zaenen et al. (2001) for a more detailed comparison of the QUD-based approaches to dialogue and monologue.

unsatisfactory answer to a question he was addressing, or in other words, that the utterance could not be grounded yet. Such a situation licenses "reraising" the same question, or alternatively, raising a question whether the utterance is true. Cooper (2003, p. 372) suggests that such question reraising corresponds to subordination structures. E.g. in (21) both sentences appear to address the question *How does John look?*, (21-b) *elaborating* on (21-a).[18] Apparently, the speaker assumes that (21-a) does not give a satisfactory account of John's appearance, and decides to dwell on the same issue in (21-b). By contrast, in (22), which exhibits a coordination structure, the first sentence is accepted as a satisfactory answer, and the speaker moves on immediately to the next question *What does John do?*

(21)   a.   John is a nice looking guy.
        b.   He has blond hair.

(22)   a.   John is a nice looking guy.
        b.   He works for a bank.

Within the class of subordinating relations, the opposition between *Elaboration* and *Explanation* could be reconstructed using van Kuppevelt's distinction (van Kuppevelt, 1995b, 1996) between *quantitatively* and *qualitatively* unsatisfactory answers (cf. Zaenen et al. (2001) for related discussion). A further development of this idea can be found in Jasinskaja (2004). The paper proposes a method to compute the semantic effects associated with the discourse relation of *Elaboration* using a QUD-based approach.

---

[18]   Note that *Elaboration* is a subordinating relation.

3.2.2   QUD-based explanations

Focus accent and focus-sensitive particles

In one of its most radical versions (e.g. Roberts, 1996), the QUD-based theory reduces the semantics of focus to its function in question-answer pairs, already discussed in Section 2.3.1. Focus introduces a *question presupposition*: the presupposed question can be simply read off the focus structure of this sentence (by replacing foci by *wh*-phrases). For instance, the sentence *Mary voted for [John]$_F$* presupposes the question *Who did Mary vote for?*; *[Mary]$_F$ voted for John* presupposes *Who voted for John?* (cf. examples (1) and (2) in Section 2.3.1); whereas a sentence with double focus *[Mary]$_F$ voted for [John]$_F$* presupposes a double question *Who voted for whom?* In other words, focus accent placement is completely determined by the underlying question in this rather straightforward way.

According to this view, the analysis of all contextual functions of focus (cf. also Sections 2.3.2 and 2.3.3) becomes a matter of DS/IS-interface. For instance, Roberts' analysis of the focus-sensitive character of *only* crucially relies on the notion of strategy and the subquestion relation between the items on the QUD stack. Roughly, a sentence with *only*, e.g. (23-b), presupposes a question "with *only*" such as (23-a). This question can in turn be successfully accommodated in the context of a corresponding question "without *only*" (24). The set of propositions associated with this question determines the alternatives that *only* quantifies over in (23-b).

(23)　　a.　Mary only invited "who" for dinner?

　　　　　b.　Mary only invited [LYN]$_F$ for dinner

(24)　　Who did Mary invite for dinner?

　　　　　a.　Mary invited John for dinner.

　　　　　b.　Mary invited Bill for dinner., etc.

The process of accommodation is governed by constraints on admissible strategies at the level of discourse structure. On the one hand, this is supposed to make sure that (23-a) can be accommodated when (24) is under discussion which renders the sequence (24)-(23-b) felicitous. On the other hand, this mechanism should provide that, for instance, the question *Who did Mary vote for?* (1-a) *cannot* be accommodated under *Who voted for John?* (1-b) to account for the fact that the sequence (1-b)-(2-a) (*Who voted for John?—Mary voted for [John]$_F$*) is infelicitous. Roberts' definition of strategy based of Groenendijk and Stokhof's notion of question entailment captures this fact.

Topic accent

Building up on Roberts' (1996) proposal, Büring (2003) uses the notions of question under discussion and strategy in order to predict the occurrence of falling A-accents ($\searrow$) and (falling-)rising B-accents ($\nearrow$) in English. He relates the former to focus [$\pm$F], and the latter to *contrastive topic* [$\pm$CT]. The basic idea is that contrastive topics, unlike foci, do not just presuppose a single question, but a whole strategy around the current utterance. Büring formalizes this idea in the definitions (25) and (26) below.[19]

(25)   CT-Congruence:
         An utterance U containing a contrastive topic can map onto a move M_U within a d-tree D only if U indicates a strategy around M_U in D.

(26) U indicates a strategy around M_U in D iff there is a non-singleton set Q' of questions such that for each $Q \in Q'$, (i) Q is identical to or a sister of the question that immediately dominates M_U, and (ii) $\|Q\|^o \in \|U\|^{ct}$

Ignoring technical details, these definitions predict, for example, that the utterance *[Fred]$_{CT}$ ate the [beans]$_F$* (where *Fred* is the contrastive topic and bears

---

[19] A *d-tree* is a discourse tree, and the term *move* is used to refer to nodes of a d-tree.

a B-accent, and *beans* is a focus and bears an A-accent) indicates a by-people strategy (cf. Section 3.2.1, pp. 180–182): *What did Fred eat?*, *What did Mary eat?*, etc. Whereas if we switch the focus and the topic marking, as in *[Fred]$_F$ ate the [beans]$_{CT}$*, the strategy indicated by this utterance is the by-food strategy: *Who ate the beans?*, *Who ate the carrots?*, etc. This analysis predicts that the CT+F and the F+CT accentuation pattern cannot be freely exchanged. For instance, (27) should be ill-formed according to this approach.

(27)   Who ate what?

      a.   [FRED ↗]$_{CT}$ ate the [BEANS ↘]$_F$.
      b.   #And [MARY ↘]$_F$ ate the [CARROTS ↗]$_{CT}$.

Further, Büring (2003) uses Schwarzschild's notion of *givenness*, which we already discussed in Sections 2.3.3 and 3.1.1 (pp. 169–170), to account for the choice between a topic accent and no accent at all. Whereas a topic accent can indicate a strategy that may include *implicit* questions (cf. Section 3.2.1, p. 178), givenness only takes into account overt moves, i.e. declarative utterances in the previous context and explicit questions. Therefore, if the subquestion in (28-b) remains implicit, the accentuation pattern in (28-d) is infelicitous, the lack of accent on *female* violates givenness. However, if (28-b), which contains the word *female*, is uttered explicitly, both (28-c) and (28-d) are possible answers.

(28)   a.   What did the pop stars wear?
      b.   (What did the female pop stars wear?)
      c.   The [FEMALE ↗]$_{CT}$ pop stars wore [CAFTANS ↘]$_F$.
      d.   #The female pop stars wore [CAFTANS ↘]$_F$.

Thus Büring's theory describes mutual constraints imposed by, on the one hand, the configuration of discourse nodes in the vicinity of an utterance and, on the other hand, topic and focus accentuation in that utterance. These constraints are

cast in terms of the notions of well-formed strategy and subquestion. However, as Büring (2003, p. 530) notes himself, in some cases a more refined rating of strategies would be necessary in order to account for the accentuation facts. He suggests that the *efficiency* of strategies could be one such measure. For instance, there is nothing wrong with the implicit question (29-b) as a subquestion of (29-a). However, going "by clothes" appears a much less efficient strategy in this case than going "by groups of pop stars", since there are normally too many different kinds of clothes, which would give rise to very long lists. Therefore, the accentuation pattern in (29-c) appears less appropriate than the pattern in (28-c) in the same (explicit) context.

(29) a.   What did the pop stars wear?

  b. #(Who wore caftans?)

  c. #The [FEMALE $\searrow$]$_F$ pop stars wore [CAFTANS $\nearrow$]$_{CT}$.

Büring does not elaborate this part of his proposal. To our knowledge, the impact of relative efficiency of strategies on discourse coherence has not yet been sufficiently investigated in connection with the contextual effects of information structure. We think, however, that this would be an important concept which would make the QUD-based theories somewhat less permissive in general, and make it possible to formulate interesting hypotheses in the domain of DS/IS-interface.

Information structure and discourse relations

Contrast: Contrast is an important notion of both discourse structure and information structure, although different things are usually meant. A recent special issue of the *Journal of Semantics* (de Hoop and de Swart, 2004) has concentrated on clearing the terminological confusion and studying the interaction of the different aspects of this notion. On the one hand, focus (contrastive fo-

cus, contrastive topic) in general involves a kind of "contrast" between the focus alternatives it induces. On the other hand, the discourse relation of contrast is said to hold between utterances that "compare" two situations with respect to their similarities and differences, or between utterances where the second contradicts a default expectation associated with the first one. In the latter case, one also often talks about *concession*, a discourse relation closely related to contrast proper and typically signaled by the same set of discourse connectives, e.g. *but* or *although*.

Umbach (2004) proposes a QUD-based analysis that integrates both the discourse-structural and the information-structural notion of contrast. First, Umbach investigates various restrictions on the sets of alternatives—the quantification domains of exhaustive operators such as *only* and bare exhaustification. Then she proposes an analysis in which the contrastive and the denial-of-expectation uses of *but* are correlated with different properties of underlying alternative sets. See also Jasinskaja (2002) and Zeevat (2004) for a more elaborate discussion and formalization of the quantification domain restrictions, as well as Kruijff-Korbayová and Webber (2000, 2001) for an account of focus sensitivity and the contrastive vs. denial-of-expectation uses of *although*.

Elaboration: Like Umbach wrt. *Contrast*, Jasinskaja (2004) uses constraints on alternative sets and the notion of exhaustification to infer the semantic effects associated with the discourse relation of *Elaboration*. By definition, *Elaboration* holds between two utterances where the first one introduces an event, and the second "elaborates" that event, i.e. adds more detail to the description of that event or some part of it. Jasinskaja does not discuss the influence of accent placement, but the general architecture of the approach is the same: the exhaustification operator quantifies over a set of alternatives constrained, on the one hand, by the current question under discussion, and on the other hand, additional constraints, such as *distinctness* (cf. Zeevat, 2004). The *Elaboration*

relation is inferred whenever two utterances address the same QUD.

Clarification: Finally, we should mention Ginzburg and colleagues' QUD-based approach to ellipsis, and various discourse relations that frequently involve ellipsis, e.g. *Correction*, *Acknowledgement* and *Clarification* (Ginzburg and Sag, 2000; Ginzburg and Cooper, 2004, Ginzburg, forthcoming). All these relations are specific to dialogue and connect utterances by distinct interlocutors. For instance, *Clarification* is a relation between an utterance by one interlocutor, and a question asked by another interlocutor, in order to clarify the content of that utterance. Two notions of the QUD framework play a central role in this analysis: the current question under discussion and the notion of grounding, cf. pp. 182–184. First, the content of an elliptical utterance is almost entirely constructed from the contextually salient question (Ginzburg and Sag, 2000). Second, Ginzburg and Cooper (2004) argue that before an utterance is sufficiently understood by all the participants of the conversation and can be *grounded*, not only the semantic objects it introduces, but also elements of its syntactic and phonological representation must be available for reference in subsequent utterances, since this kind of reference is necessary for a proper analysis of *Clarification* questions. In accordance with these insights, Ginzburg and Cooper account for the ambiguous character of interrogatives like *Finagled?* (in the context of *Did Bo finagle a raise?*) which can be paraphrased either as *Are you asking if Bo (of all actions)* FINAGLED *a raise?*, or as *What does it mean to finagle?*

3.3 Discussion

We do not intend to engage in a full scale comparison of the predictions of the theoretical approaches presented above. Instead, we would like to summarize the findings of previous studies and formulate our tentative answer to the

question of which aspects of discourse structure interact with which aspects of information structure, as far as their pragmatic interaction is concerned.

In Section 2.1, we articulated three types of information that can be encoded in discourse structure: (1) discourse constituency, (2) discourse coordination vs. subordination, and (3) specific semantically/pragmatically motivated discourse relations, such as *Elaboration, Contrast, Narration*, etc.

On the basis of the findings in the literature, one could claim that information structure, as manifested by accent choice and placement, stands in especially close interaction with the third of these elements—the choice of a discourse relation. That is, once we know (or hypothesize) that the utterances u_1 and u_2 form a discourse constituent, the information structure of these utterances helps us constrain the choice among possible discourse relations. The most uncontroversial is the connection between the discourse relation of *Contrast* and contrastive topic/focus marking. But *Concession*, as well as various dialogue-specific relations, also appears to be affected by information structure, mainly by the location of the focus accent. At the same time, the impact of other discourse relations (e.g. *Narration*) on information structure is not so well-studied and awaits further research.

On the other hand, discourse relations are also the most controversial part of discourse structure, and some theories e.g. Grosz and Sidner (1986) or the QUD-approaches do not acknowledge them. In this group of theories the relation between information structure and the choice of a semantic/pragmatic relation between utterances is not direct, but is mediated by speakers' intentions (e.g. as QUDs). That is, information structure helps the hearer recover the set of issues the speaker intends to discuss, which in turn affects the semantic relations between the individual utterances. However, up to now there has been no extensive theoretical proposal that accounts for the connection between questions under discussion and the semantic relations between declarative utterances. This is again a desideratum for the future.

But as was mentioned, the choice of a discourse relation between two utterances presupposes the existence of a structural connection between these utterances (they have to form a discourse constituent). So what about discourse constituency, and other "properly structural" aspects of discourse structure? Do they have anything to do with information structure and accenting or not? Sporadic mentions in the literature (e.g. Schwarzschild, 1999, pp. 165–166) suggest that they do, and Nakatani (1997) confirmed it at least for pronouns. In brief, the presence vs. absence of a structural relation, as well as coordinating vs. subordinating character of the relation determines which individuals, properties, situations etc. are *salient* at the current point in discourse, and that in turn affects which items can be considered as given or new, and hence whether they should be accented or uttered without an accent. This is another field in the domain of pragmatic IS/DS-interface which calls for further investigation.

4 DS and IS Interact at the Level of Phonetics

As previously mentioned, prosodic correlates of information structure and discourse structure exploit the same phonetic parameters. IS-driven prosodic phenomena like pitch accent position and type, deaccentuation, short pauses in the vicinity of focus exponents, or decreased speaking rate on focus exponents are mostly local phenomena operating on words or syllables. DS-driven phenomena, on the other hand, tend to be global, i.e. affecting whole phrases (pitch range, general speaking rate and intensity) or the strength of separation of adjacent phrases (pause duration and type). But this distinction along the local/global dimension is not complete. There is a gray area where IS- and DS-driven prosodic phenomena come very close and may impose conflicting requirements. According to this situation, we distinguish two types of interaction on the phonetic/phonological level: regular interactions and instances of conflicting requirements.

Conflicting requirements can be expected particularly with pauses and speaking rate. Both parameters are used to signal IS- as well as DS-driven prosodic phenomena. Information structure is one source of prosodic phrasing which is implemented—among other features—by pauses between phrases and by phrase-final decrease of speaking rate ('final lengthening'). Discourse prosody, on the other hand, usually falls back on the existing phrasing and controls pause duration and more general variations of speaking rate, i.e. rate differences between phrases, not within phrases. But besides the dividing function, pauses and speaking rate variation may also be employed by IS to isolate and highlight focused items, as mentioned in Section 2.4. This leads to a revision of the neutral, syntax-based phrasing and potentially entails conflicts with DS-derived phrasing requirements, since discourse structure is crucially linked to syntactic structure.

Regular interactions are due to the similarity of the expressive phonetic parameters. In these cases, both IS- and DS-requirements are fulfilled, but with noticeable modifications in one area depending on specific parameter settings in the other area. An interesting example of this type of interaction is the influence of compressed pitch range on the realization of pre-nuclear pitch accents. In phrases with an overall falling contour, the first accent is usually realized with the highest pitch peak, while following accents are downstepped, resulting in the lowest pitch-peak assigned to the nuclear (i.e. last) accent (see Figure 4, top). In phrases with compressed pitch range this mechanism would lead to very low pitch peaks on nuclear accents (see Figure 4, bottom left), further disguised by creaky voice which often co-occurs with compressed pitch range. It was never systematically investigated whether this really happens or whether other strategies are available to the speaker to preserve an appropriate prominence level of the nuclear accent. This question is of particular interest to the discussion on the theoretical status of discourse prosody. Assuming the standard view that discourse related variation of global prosodic parameters is not part

of the phonological representation of an utterance but directly implemented at the phonetic level, the phonetic process of gradual lowering of pitch accents is the only possibility of the production system to adapt local pitch events to the compressed range. If we assume, however, that discourse prosody is phonologically represented, then local and global prosodic aspects of an utterance can interact at the phonological level, which allows at least two additional strategies to handle the problem at hand. If the phonological system is aware of the compressed pitch range one possibility would be to inhibit the phonological process of downstep, resulting in a nuclear accent with equal height to the prenuclear accents. This, however, is not likely because it would possibly change the metrical pattern and subsequently the interpretation of the utterance (Ladd, 1996). We assume that a deaccentuation strategy is more likely (see Figure 4, bottom/right). Deaccentuation of prenuclear pitch accents leaves the full range to the realization of the nuclear accent, while the metrical strength of the prenuclear items can be maintained using intensity or durational prominence features (cf. Beaver et al. (2002) for the realization of second occurrence focus). An investigation of this problem has to be based on production data as well as perception experiments, since particularly the perceptual function of non-pitch prosodic features regarding sentence-level prominence is still unclear.

5 Conclusions

In this paper we presented the current state of research on the issue of how discourse structure and information structure interact. We propose that this interaction should be investigated at (at least) two levels. First, these two structures impose mutual constraints at the pragmatic level. This aspect of DS/IS interaction has been in the focus of linguists' interest for a some time now, and we have tried to give a comprehensive survey of relevant studies. However, we have found a lot of theoretical and empirical issues that still have to be clarified in
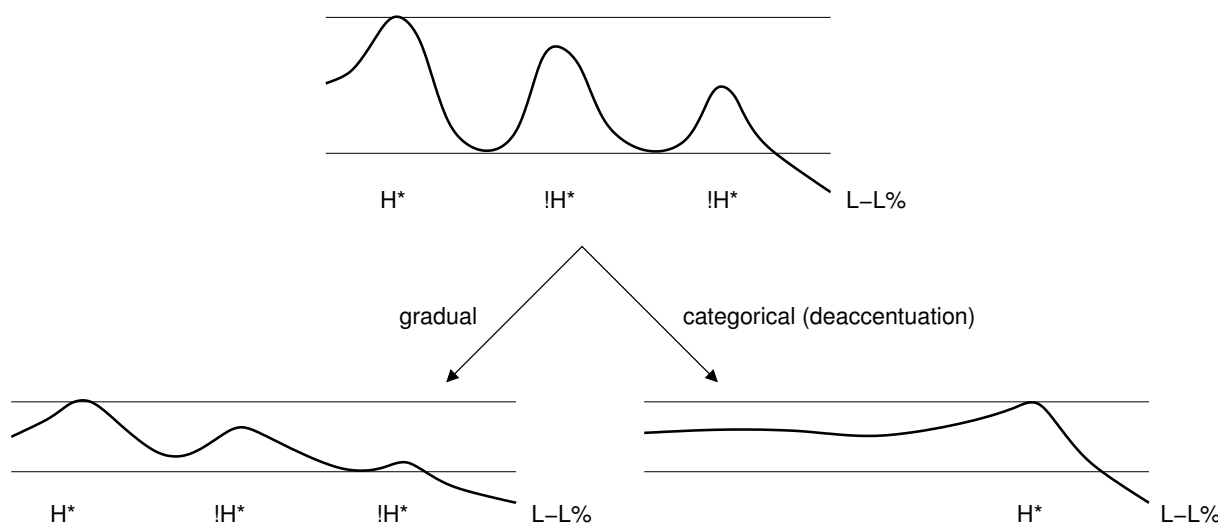
Figure 4: Phonetic (bottom/left) and phonological (bottom/right) adaption of multi-accent utterances to compressed pitch range.

order to complete the picture.

Empirical testing is particularly called for by theories that treat the contextual effects of IS as a species of anaphora, such as Schwarzschild's (1999) theory of givenness, cf. Section 3.1.1, pp. 169–170. Combined with a theory of anaphoric accessibility (based on one of the available discourse models, e.g. DRT, SDRT, or LDM), an anaphoric approach to IS provides clear testable hypotheses about the influence of discourse constituency and subordination relations on accent placement. Verifying these hypotheses appears to be a highly relevant and realistic research task.

On the theoretical side, the QUD-based approach provides a vast field for further work. As we have seen in Section 3.2, the whole architecture of this approach is such that it "postpones" the treatment of most IS phenomena to the level of discourse structure. Contextual effects that are traditionally treated at the level of focus semantics are now cast in terms of constraints on accommodation of (implicit) questions, or more generally, constraints on possible inquiry strategies. But the solutions proposed so far are not sufficient. One possible di-

rection for strengthening these theories is to rate strategies according to their efficiency, as suggested by Büring (2003, p. 530). In general, it should be said that if the QUD framework's ambition is to provide a general theory of discourse structure and semantics, a number of rather central components still have to be developed. For instance, comprehensive accounts of discourse anaphora, as well as semantic relations between utterances, are still missing. Once these missing parts are supplied, the framework is likely to provide a rather broad notion of DS/IS interface and at the same time a uniform treatment of various interface phenomena.

Finally, the interaction of prosodic topic/focus marking with the choice of semantic discourse relation leaves space for both theoretical and empirical investigations. So far, the only discourse relations whose information-structural effects have been studied reasonably well are *Contrast*, *Correction* and *Question-Answer Pair*, cf. Section 3.1.1, pp. 170–172 and Section 3.2.2, pp. 189–190. It is still unclear whether other relations, such as *Narration*, *Elaboration*, *Explanation*, have their characteristic information structures and accentuation patterns. Some interesting ideas on this issue were sketched out by Asher and Txurruka (1995) and need to be both worked out theoretically and tested empirically.

These research areas are suggested by the current literature and they all pertain to the study of DS/IS interface at the level of pragmatics. However, as we emphasized in this paper, discourse structure and information structure address a number of prosodic devices which in turn interact at the level of phonetics (or phonology). These phenomena have received much less attention in general and have never been brought up in connection with the issue of DS/IS interface. We have outlined some directions in which this line of research could go.

Bibliography

Arnold, J., Fagnano, M., and Tanenhaus, M. (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research*, 32:25–36.

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. Studies in Linguistics and Philosophy. Kluwer Academic Publisher, Dordrecht.

Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Studies in Natural Language Processing. Cambridge University Press.

Asher, N. and Txurruka, I. G. (1995). Extending SDRT to integrate a theory of the informational partition. Technical Report ILCLI-95-LIC-3, Institute for Logic, Cognition, Language and Information, Euskal Herriko Universidad.

Asher, N. and Vieu, L. (2003). Subordinating and coordinating discourse relations. *Lingua, to appear*.

Ayers, G. M. (1994). Discourse functions of pitch range in spontaneous and read speech. *Ohio State University Working Papers in Linguistics*, 44:1–49.

Beaver, D., Clark, B., Flemming, E., and Wolters, M. (2002). Second occurrence focus is prosodically marked. Ms., Stanford University.

Büring, D. (1997). The great scope inversion conspiracy. *Linguistics and Philosophy*, 20:175–194.

Büring, D. (2003). On D-trees, beans and B-accents. *Linguistics and Philosophy*, 26:511–545.

Chafe, W. (1976). Giveness, contrastiveness, definiteness, subject, topics, and point of view. In Li, C., editor, *Subject and Topic*, pages 27–55. Academic Press, New York.

Cooper, R. (2003). Afterword (to the special issue on discourse and information structure). *Journal of Logic, Language and Information*, 12:369–374.

Cristea, D., Ide, N., and Romary, L. (1998). Veins theory: A model of global discourse cohesion and coherence. In *Proceedings of COLING-ACL 1998*, pages 281–285.

Culicover, P. W. and McNally, L., editors (1998). *The Limits of Syntax*, volume 29 of *Syntax and Semantics*. Academic Press.

de Hoop, H. and de Swart, P. (2004). Contrast in discourse: Guest editors' introduction. *Journal of Semantics*, 21:87–93.

den Ouden, H., Noordman, L., and Terken, J. (2002). The prosodic realization of organizational features of texts. In *Proceedings of Speech Prosody 2002*, pages 543–546, Aix-en-Provence.

Frota, S. (2000). *Prosody and Focus in European Portuguese. Phonological Phrasing and Intonation.* Garland, New York.

Gardent, C. (1997). Discourse tree adjoining grammar. Claus report 89, University of the Saarland, Saarbrücken.

Ginzburg, J. (1995). Resolving questions 1 & 2. *Linguistics and Philosophy*, 18:459–527, 567–609.

Ginzburg, J. (1996a). Dynamics and the semantics of dialogue. In Seligman, J. and Westerståhl, D., editors, *Logic, Language and Computation*, volume 1. CSLI Publications, Stanford, CA.

Ginzburg, J. (1996b). Interrogatives: Questions, facts and dialogue. In Lappin, S., editor, *The Handbook of Contemporary Semantic Theory*. Blackwell, Oxford.

Ginzburg, J. (1997). On some semantic consequences of turn taking. In Dekker, P., Stokhof, M., and Venema, Y., editors, *Proceedings of the 11th Amsterdam Colloquium*, ILLC-Publications, Amsterdam.

Ginzburg, J. (1999a). Fragmenting meaning: Clarification ellipsis and nominal anaphora. In Bunt, H., editor, *Proceedings of the 3rd Workshop on computational semantics*, Tilburg, The Netherlands.

Ginzburg, J. (1999b). Semantically-based ellipsis resolution with syntactic presuppositions. In Bunt, H. and Muskens, R., editors, *Computing Meaning: Current issues in computational semantics*. Kluwer, Dordrecht.

Ginzburg, J. (forthcoming). *Semantics and Interaction in Dialogue*. CSLI Publications and University of Chicago Press, Stanford CA. Draft chapters available from `http://www.dcs.kcl.ac.uk/staff/ginzburg`.

Ginzburg, J. and Cooper, R. (2004). Clarification, ellipsis, and the nature of contextual updates in dialogue. *Linguistics and Philosophy*, 27:297–365.

Ginzburg, J. and Sag, I. A. (2000). *Interrogative Investigations: The From, Meaning and Use of English Interrogatives*. CSLI Publications, Stanford CA.

Groenendijk, J. and Stokhof, M. (1984). *Studies in the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam.

Grosz, B. and Hirschberg, J. (1992). Some intonational characteristics of discourse structure. In *Proceedings of the 2nd International Conference on Spoken Language Processing*, pages 429–432, Banff, Canada.

Grosz, B., Joshi, A. K., and Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.

Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Grüning, A. and Kibrik, A. A. (2002). Referential choice and activation factors: A neural network approach. 4th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC'2002). Lisbon, Portugal.

Gustafson-Capková, S. and Megyesi, B. (2002). Silence and discourse context in read speech and dialogues in Swedish. In *Proceedings of Speech Prosody 2002*, pages 363–366, Aix-en-Provence.

Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10:41–53.

Hendriks, P. (2004). Coherence relations, ellipsis and contrastive topics. *Journal of Semantics*, 21:133–153.

Herman, R. (2000). Phonetic markers of global discourse structures in English. *Journal of Phonetics*, 28:466–493.

Higginbotham, J. and May, R. (1981). Questions, quantifiers and crossing. *Linguistic Review*, 1:41–79.

Hirschberg, J. and Nakatani, C. H. (1996). A prosodic analysis of discourse segments in direction-giving monologues. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*, pages 286–293, Santa Cruz.

Horne, M., Frid, J., and Roll, M. (2004). Hesitation disfluencies after the clause marker *att* 'that' in Swedish. *Lund University Working Papers in Linguistics*, 51.

Horne, M., Hansson, P., Bruce, G., and Frid, J. (1999). Prosodic correlates of infomation structure in Swedish human-human dialogues. In *Proceedings of EURO-SPEECH99*, pages 29–32, Budapest.

Jasinskaja, E. (2004). Exhaustification and semantic relations in discourse. In Geurts, B. and van der Sandt, R., editors, *Proceedings of the ESSLLI'2004 Workshop on Implicature and Conversational Meaning*, pages 14–19, Nancy, France.

Jasinskaja, K. (2002). Relevance and other constraints on the quantification domain of *only*. In *Proceedings of the Workshop on Information Structure in Context*. IMS Stuttgart.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer Academic Publishers.

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, 1:3–44.

Klein, W. and von Stutterheim, C. (1987). Quaestio und die referentielle Bewegung in Erzälungen. *Linguistische Berichte*, 109:163–185.

Klein, W. and von Stutterheim, C. (1992). Textstruktur und referentielle Bewegung. *Zeitschrift für Literaturwissenschaft und Linguistik*, 86:67–92.

Klein, W. and von Stutterheim, C. (to appear, 2001). How to solve a complex verbal task: Text structure, referential movement and the quaestio. In Espinas, G. S. and Brum, M., editors, *Recherche sur l'acquisition des langues*.

Kodzasov, S. V. (1996). Zakony frazovoj akcentuacii. In Nikolaeva, T. M., editor, *Prosodičeskij stroj russkoj reči*, pages 181–204. Institut russkogo jazyka Rossijskoj akademii nauk, Moskva.

Komogata, N. (2003). Information structure in subordinate and subordinate-like clauses. *Journal of Logic, Language and Information*, 12:301–318.

Kruijff-Korbayová, I. and Steedman, M. (2003). Discourse and information structure. *Journal of Logic, Language and Information*, 12:249–259.

Kruijff-Korbayová, I. and Webber, B. L. (2000). Discourse connectives, inference and information structure. In Bos, J. and Kohlhase, M., editors, *Workshop on Inference in Computational Semantics (ICoS-2)*, pages 105–120, Schloß Dagstuhl, Germany. International Conference and Research Center for Computer Science.

Kruijff-Korbayová, I. and Webber, B. L. (2001). Concession, implicature, and alternative sets. In Bunt, H., editor, *4th International Workshop on Computational Semantics (IWCS-4)*, pages 227–248, Tilburg, The Netherlands.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge University Press.

Larsson, S. (2002). *Issue-based Dialogue Management*. PhD thesis, Göteborg University.

Lehiste, I. (1975). The phonetic structure of paragraphs. In Cohen, A. and Nooteboom, S. G., editors, *Structure and Process in Speech Perception*, pages 195–203. Springer.

Lehiste, I. (1979). Perception of sentence and paragraph boundaries. In Lindblom, B. and Öhman, S., editors, *Frontiers of Speech Research*, pages 191–201. Academic Press, London.

Mann, W. C. and Thompson, S. (1988). Rethorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

Mayer, J. (1999). Prosodische Merkmale der Diskursrelationen. *Linguistische Berichte*, 177:65–86.

Möhler, G. and Mayer, J. (2001). A discourse model for pitch-range control. In *Proceedings of the 4th ISCA Tutorial and Research Workshop on Speech Synthesis*, Perthshire, Scotland.

Nakajima, S. and Allen, J. F. (1993). A study on prosody and discourse structure in cooperative dialogues. *Phonetica*, 50:197–210.

Nakatani, C. H. (1997). Discourse structural constraints on accent in narrative. In van Santen, J., Sproat, R. W., Olive, J., and Hirschberg, J., editors, *Progress in Speech Synthesis*, pages 139–156. Springer, New York.

Partee, B. H. (1996). Allegation and local accomodation. In Partee, B. H. and Sgall, P., editors, *Discourse and Meaning: Papers in Honor of Eva Hajičová*, pages 65–86. John Benjamins, Amsterdam.

Passonneau, R. J. and Litman, D. J. (1993). Feasibility of automated discourse segmentation. In *Proceedings of the 31th Annual Meeting of the Association for Computational Linguistics*, pages 148–155, Ohio.

Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. PhD thesis, MIT.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, 12:601–638.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics*, 49:91–136.

Roberts, C. (1998). Focus, the flow of information, and universal grammar. In Culicover and McNally (1998), pages 109–160.

Rooth, M. E. (1985). *Association with Focus*. PhD thesis, University of Massachusets, Amherst.

Rooth, M. E. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1:75–116.

Schlangen, D. (2003). *A Coherence-Based Approach to the Interpretation of Non-Sentential Utterances in Dialogue*. PhD thesis, School of Informatics, University of Edinburgh, Edinburgh, UK.

Schlangen, D. (2004). Causes and strategies for requesting clarification in dialogue. In *Proceedings of the 5th Workshop of the ACL SIG on Discourse and Dialogue*, Boston, USA.

Schwarzschild, R. (1999). GIVENness, AVOIDF and other constraints on the placement of accent. *Natural Language Semantics*, pages 141–177.

Selkirk, E. (1995). Sentence prosody: Intonation, stress and phrasing. In Goldsmith, J., editor, *Handbook of Phonological Theory*, pages 550–569. Blackwell, Oxford, UK.

Sluijter, A. M. C. and Terken, J. M. B. (1993). Beyond sentence prosody: Paragraph intonation in Dutch. *Phonetica*, 50:180–188.

Smith, C. L., McCraw, M. O., and Hogan, L. A. (2002). The role of topic structure in speech timing: A comparison of English and Japanese. Talk at LabPhon VIII.

Swerts, M. (1997). Prosodic features at discourse boundaries of different strength. *Journal of the Acoustical Society of America*, 101:514–521.

Swerts, M., Avesani, C., and Krahmer, E. (1999). Reaccentuation or deaccentuation: A comparative study of Italian and Dutch. In *Proceedings of the XIVth International Congress of Phonetic Science*, pages 1541–1544, San Francisco.

Swerts, M. and Geluykens, R. (1993). The prosody of information units in spontaneous monologue. *Phonetica*, 50:189–196.

Swerts, M. and Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech*, 37(1):21–43.

Taboada, M. (2002). Centering and pronominal reference: In dialogue, in Spanish. In *Proceedings of the 6th Workshop on the Semantics and Pragmatics of Dialogue (EDILOG)*, Edinburgh, UK.

Umbach, C. (2004). On the notion of contrast in infromation structure and discourse structure. *Journal of Semantics*, 21:155–175.

Vallduví, E. (1992). *The Informational Component*. Garland, New York.

Vallduví, E. and Vilkuna, M. (1998). On rheme and kontrast. In Culicover and McNally (1998), pages 79–108.

van Donzel, M. (1999). *Prosodic Aspects of Information Structure in Discourse*. PhD thesis, The Hague University.

van Kuppevelt, J. (1995a). Discourse structure, topicality and questioning. *Journal of Linguistics*, 31:109–147.

van Kuppevelt, J. (1995b). Main structure and side structure in discourse. *Linguistics*, 33:809–833.

van Kuppevelt, J. (1996). Directionality in discourse: Prominence differences in subordination relations. *Journal of Semantics*, 13:363–395.

van Kuppevelt, J. (1999). Topic and comment. In Brown, K. and Miller, J., editors, *Concise Encyclopedia of Grammatical Categories*, pages 369–374. Elsevier, Amsterdam.

van Rooy, R. (2003). Questioning to resolve decision problems. *Linguistics and Philosophy*, 26:727–763.

Venditti, J. and Swerts, M. (1996). Prosodic cues to discourse structure in Japanese. In *Proceedings of the International Conference on Spoken Language Processing*, pages 725–728, Philadelphia.

von Stutterheim, C. and Klein, W. (1989). Referential movement in descriptive and narrative discourse. In Dietrich, R. and Graumann, C. F., editors, *Language Processing in Social Context*, pages 39–76. North Holland, Amsterdam.

Webber, B., Joshi, A., Stone, M., and Knott, A. (2003). Anaphora and discourse structure. *Computational Linguistics*.

Zaenen, A., Larsson, S., Sandor, A., and Traum, D. (2001). Text and dialogue. In *The Trindi Book*, chapter 9, pages 261–296.

Zeevat, H. (2004). Contrastors. *Journal of Semantics*, 21:95–112.

Ekaterina Jasinskaja, Jörg Mayer and David Schlangen
Universität Potsdam
SFB 632, Institut für Linguistik
Postfach 601553
14415 Potsdam
Germany
jasinsk@rz.uni-potsdam.de
{mayer,das}@ling.uni-potsdam.de

Focus Strategies in Chadic: The Case of Tangale Revisited[*]

Katharina Hartmann and Malte Zimmermann

Humboldt University Berlin

We argue that the standard focus theories reach their limits when confronted with the focus systems of the Chadic languages. The backbone of the standard focus theories consists of two assumptions, both called into question by the languages under consideration. Firstly, it is standardly assumed that focus is generally marked by stress. The Chadic languages, however, exhibit a variety of different devices for focus marking. Secondly, it is assumed that focus is always marked. In Tangale, at least, focus is not marked consistently on all types of constituents. The paper offers two possible solutions to this dilemma.

Keywords: tone languages, focus marking, focus movement

1 Introduction

This paper investigates the focus systems of some Chadic languages, in particular Tangale, a Western Chadic language spoken in the North of Nigeria. We show that standard focus theories, which are based on stress languages, cannot account for the rich variety of focus phenomena found in the Chadic tone languages. The standard theories assume that focus is obligatorily marked by stress. The Chadic languages, however, choose from a variety of devices for

---

focus marking (movement, morphological marking, prosodic phrasing). Apart from this, the formal means of focus marking sometimes depend on the syntactic category of the focus constituent, which can result in a systematic underspecification of focus: at least in Tangale, it appears that focus is not marked consistently on all types of constituents. The data discussed suggest that universal theories of focus have to be either more complex than so far assumed. Or, they could still be simple, but would have to allow for a certain degree of underspecification in focus marking. The second alternative shifts much of the interpretive burden to the pragmatic component.

In section two, we give a definition of focus and present our view of the standard theory. In section three, we show that tone languages sometimes use more than just one strategy to mark a focus, thereby deviating from one of the core assumptions of the standard theory. In sections four and five, we concentrate on the tonal languages of the Chadic family, especially on Tangale, a language spoken in Northern Nigeria (Gombe State). Our investigation shows that at least some Chadic languages seem to have more than one focus marking device at their disposal, suggesting a modification of the standard theory. However, in section six, we go on to show that only subjects are consistently marked for focus in Tangale. In contrast, focus on all other constituents is only sporadically marked and must therefore be heavily supported by the pragmatic system. This might bring us back to the assumption of the standard theory, that there is only one mechanism of (obligatory) focus marking.

2   Standard Focus Theories

2.1 A definition of focus

We adopt the following definition of focus for tone and stress languages (which is independent of focus marking): Focus on a constituent α ($[\alpha]_F$) invokes a set A of alternatives to α, indicating that members of A are under consideration (Rooth 1985). Depending on the interaction of α with other alternatives, a focus can receive different pragmatic readings: A focus is *corrective* if α replaces an element of A previously introduced into the common ground (CG), i.e. the linguistic context preceding $\forall$, see (1a). A focus is *selective* if α introduces an element of A into the CG and some elements of A are made explicit, see (1b). A focus expresses *new-information* if α introduces an element of A into the CG and the members of A are left implicit, see (1c).

(1)   a.   (Peter painted his bicycle red.) No, he painted it $[blue]_F$.
            α = blue, A = {blue, red, green, pink,…}

      b.   (Did Peter paint his bicycle red or blue?) He painted it $[blue]_F$.
            α = blue, A = {blue, red, green, pink,…}

      c.   (Which color did Peter paint his bicycle?) He painted it $[blue]_F$.
            α = blue, A = {blue, red, green, pink,…}

The alternative sets in (1a-c) are identical. This shows that the foci do not differ semantically, but only pragmatically in the sense illustrated above. In our view, focus as defined above is a universal category. The focus marking devices, however, vary considerably across the world's languages. This paper investigates means of focus marking in Chadic tone languages and compares them to focus marking in stress languages.

2.2 Standard focus theories

Theories of focus are usually based on the properties of stress languages. This bias towards a certain typological kind of languages has accompanied the development of focus theories since focus became a subject of scientific interest. In a nutshell, standard focus theories make the following three assumptions: Firstly, focus must be marked. Secondly, there is a single strategy to mark a focus, which is stress. And thirdly, any syntactic category can be focused.

Jackendoff (1972) already states that: "If a phrase P is chosen as the focus of the sentence S, the highest stress in S will be on the syllable of P that is assigned highest stress by the regular stress rules." (p. 237). Following Jackendoff, the relation between the (pragmatically determined) size of a focus and placement of stress is mediated by a syntactic focus (F-) feature. The F-feature marks the focus of a sentence. The stress must be realised within the F-marked constituent (Jackendoff 1972:240f).

F-features also form the back bone of Selkirk's focus theory (Selkirk 1984, 1995). In this approach, F-feature assignment is not primarily triggered by pragmatics, but by phonetic conditions: the constituent carrying main stress receives an F-feature (the *Basic Focus Rule*, Selkirk 1995:555). This feature can project along the functor-argument structure. If the accented constituent is a complement, it projects to the selecting head. If it is a head, it projects to the head's maximal projection (*Focus Projection*, Selkirk 1995:555). Focus projection enables a constituent which is bigger than the stress bearing unit to be the focus of a sentence. Constituents which are F-marked (and are not the sentence focus) are interpreted as new in the discourse (Selkirk 1995:556). The following examples illustrate the working of Selkirk's theory. The stress bearing constituent is printed in capitals.

(2)   a.   What did Carolin bring to the party?
          She brought [$_{NP}$ SALAD]$_F$

      b.   What did Carolin do?
          She [$_{VP}$ brought$_F$ [SALAD]$_F$]$_F$

In (2a), the accented object is F-marked. It is the focus of the sentence since it replaces the *wh*-word of the question. In (2b) the *wh*-question requires a predicate focus. Again, the accented object receives an F-feature, which projects across V to VP, thereby defining the focus of the sentence.

Schwarzschild (1999) and Büring (2004) examine the validity of focus projection rules showing that these rules are empirically inadequate. The heart of their argument is that any accent within an XP can project focus given an appropriate context. Thus, F-marking of XP does not require an accent on X^0 or on the complement of X^0. This is shown in (3), taken from Büring (2004:7). This example shows that focus can project from unergative subjects, which is excluded in Selkirk's theory since the subject is neither a complement nor a head.

(3)   Q: Why did Helen buy bananas?

      A: [Because JOHN bought bananas]$_F$

Büring and Schwarzschild maintain the assumption that a focus must be maximally prominent and that it must be marked somewhere within the focused constituent (see also Truckenbrodt 1995, 1999). The position of the main stress thus depends on the argument structure in a less direct way then hitherto proposed.

To summarise, the standard theories assume that focus on any constituent is marked by one and the same strategy. The only factor to be considered is

stress. Additional means of highlighting a focus constituent (i.e. clefting or movement) are possible in stress languages, but they are always accompanied by an accent on the clefted/moved constituent, as shown by the following example.

(4) a.   A BOOK, Peter bought (not a REcord).

     b.   It is a BOOK that Peter bought (not a REcord).

3   Focus in tone languages

The assumption that focus is marked by only one factor does not hold for all tone languages. This is illustrated by two examples: In Mandarin Chinese, focus is indicated by two factors, movement and stress (manifested as length and intensity): Focused constituents which do not appear in their (sentence final) default position are likely to be stressed. Postfocal material is destressed (see Xu 1999, Xu 2004; the data in (5) are from Xu 2004:291).

(5) a.   Shui   lai-le?
          who   came
          'Who has come?'

     b.   Lai-le [jige   meiguoren]$_F$ *(focus default position)*
          came   some Americans
          'Some Americans have come.'

     c.   [Jige MEIGUOREN]$_F$ lai-le   *(non-default position)*

In Tupuri (Niger-Congo) focus is sometimes indicated by an ex situ (cleft) strategy (6a), from Ruelland (2000), and sometimes by reduplication (6b), our data (unfortunately without tones).

(6) a. t$_1$ wɔ̄ ɗè pṹy tí dárgè dìŋ táktíbáy$_1$
       go with hyena to hunt COP bat
       'It is Bat that will go hunting with Hyena.'

   b. A juujuu gi, a ri súu ga
       he drink-drink ?? he ate yesterday NEG
       'He DRANK, but he didn't eat yesterday.'

The data in (5) and (6) show that at least some tonal languages exhibit more than one focus strategy. While in Chinese the choice of strategies seems to depend on structural factors (focused constituent sentence final or not), in Tupuri the choice of strategy depends on the syntactic category (focused constituent verbal or not).

4 Focus in Chadic Languages

4.1 DP-focus in Chadic

Focusing of DP-arguments is well-documented for Hausa (see Newman 2000, Jaggar 2001) and for a range of other Chadic languages (see Tuller 1987, 1992, Frajzyngier 1989, 1993, 2001, 2002, Schuh 1998, 2004).

4.1.1 Focus movement

A common strategy of focusing a DP-constituent in Chadic is to move it to a designated position. Often, the resulting structure has a cleft-like nature and a lexical focus marker (in many cases formally identical to the copula or the relative marker). Movement may also be accompanied by high tone raising of the fronted constituent (Hausa, see Leben et al. 1989), or by a change in verbal aspect (Hdi, Frajzyngier 2002). Focus movement can target several positions, namely to the sentence-initial position, to a postverbal position, or to the sentence-final position. We will consider each kind in turn.

In Hausa, an SVO language, focused DPs are fronted to the sentence-initial position (cf. Newman 2000). After the fronted constituent, a focus marker (FOC) is optionally inserted. (7a) is an example with neutral (i.e. all new) focus. In (7b), the object is focused and appears sentence-initially.

(7) a. Bintà zaa tà   biyaa teelà   *(neutral)*
        B.   FUT 3sg.f pay   tailor
        'Binta will pay the tailor.'

    b. *teelà₁ (nee)* Bintà zaa tà   biyaa t₁   *(OBJ-focus)*
        tailor FOC   B.   FUT 3sg.f pay
        'Binta will pay the TAILOR.'

Focus fronting also occurs in Hdi, a VSO language documented in Frajzyngier (2002). (8a) is a neutral example again. In (8b), the focused object is fronted. In addition to fronting, there is a change in verbal aspect (see Frajzyngier 2002:408; SO = point of view of reference, REF = referential, SEQ = sequential marker).

(8) a. kà ks-ú-tá   ùvá tá vàzák   *(neutral)*
        SEQ touch-SO-REF cat OBJ rooster
        'And Cat devoured Rooster.'

    b. [ghùz-á xìyá]₁   yà tà sə mbítsá t₁
        beer-GEN guinea corn DEM IMPF drink M.
        'It is the corn beer that Mbitsa drinks.'   *(OBJ-focus)*

Focused constituents are also fronted to the sentence-initial position in Kanakuru (Tuller 1992) and Pero (Frajzyngier 1989).

The second strategy of focus movement observed in the Chadic languages is movement to a postverbal position. For an illustration of this strategy, consider the following Tangale data (from Kidda 1993:30f; due to the

phonological process of vowel deletion, cf. also section 5.1, the name *Laku* sometimes appears as *Lak*).

(9) a. Lak padu-g   landá   *(neutral)*
        L.   buy-PERF dress
        'Laku bought a dress.'

    b. padu-g   landá nóŋ tom tíjo?   *(SUBJ-focus)*
        buy-PERF dress who from T.
        'Who bought a dress from Tijo?'

Tangale is an SVO language; (9a) represents the neutral word order. If a subject is focused as in (9b) (a *wh*-focus), it is obligatorily displaced from its initial base position to a postverbal position. The Tangale focus system will be discussed in detail in section 5. Focus movement to a postverbal position also takes place in Bade, Podoko, Kanakuru, and Ngizim (cf. Tuller 1992).

Focused constituents can also appear in sentence-final position, as evidenced by the following example from Ngizim (SVO, Tuller 1992). In (10), the subject is focused, it consequently appears in sentence-final position. This strategy is also testified in Tangale (Tuller 1992), Bole (Schuh 2004, cf. also footnote 5), and Pero (Frajzyngier 1989).

(10)   ɗəbdə karee aa aasək   nən Audu   *(SUBJ-focus)*
       sold   goods in market FOC A.
       'AUDU sold the goods in the market.'

4.1.2 In situ focus

In some languages, focused DPs remain in situ. In this case, prominence is achieved by morphological, aspectual, or prosodic marking. Consider the Mupun examples in (11) (from Frajzyngier 1993). The focused object DP is not

displaced from its base-generated position (Mupun is an SVO language). Focus is only indicated by the presence of the focus marker *a*.

(11)   war cet a   lua ba a   pupwap kas.   *(OBJ-focus)*
        3f cook FOC meat NEG FOC fish   NEG
        'She cooked MEAT, not FISH.'

In Miya (Schuh 1998), the verbal aspect changes in order to indicate focus. In (12b), the object is focused. The aspectual change is manifested in the absence of the discontinuous totality marker (TOT) *suw…ay*, which is present in the neutral example (12a).

(12) a.   à   már suw zhàak-áy   *(neutral)*
           he got TOT donkey-TOT
           'He got a donkey.'

     b.   à   már zhàakə   *(OBJ-focus)*
           he got donkey
           'He got a DONKEY.'

In situ focus is also possible in Lele, where it is indicated by a focus marker (see Frajzyngier 2001). In Pero, in situ focus is marked by an intonational break before the focused element (cf. Frajzyngier 1989). Focus constituents can also remain in situ in Ga'anda (cf. Ma Newman 1971) and in Hausa, where it is not evident if and how in situ foci are marked (cf. Jaggar 2001 and Green and Jaggar 2002).

To sum up, the Chadic languages express focus on DP-arguments by using different markers of prominence. DP-focus is indicated by movement (Hausa, Hdi, Tangale, Kanakuru, Ngizim, Bade, Bole, Pero), by morphological marking (Mupun, Lele), by changes in the verbal aspect (Miya), or by different prosodic

phrasing (Pero). Languages that mark focus by movement sometimes use morphological marking or a change of verbal aspect in addition. Their grammatical systems appear to be somewhat uneconomical with respect to focus marking.

With the exception of Pero (focus fronting and prosodic phrasing) and Hausa (focus fronting and in situ focus), the Chadic languages discussed here employ a single strategy to mark DP-focus. This suggests the following preliminary hypothesis:

(13)   *Preliminary Hypothesis (to be refuted):*
         In general, Chadic languages employ only a single focus strategy.

We will see below that this hypothesis cannot be maintained on closer inspection.

4.2 V(P)-focus: The picture changes

Concerning the realisation of predicate focus, the Chadic languages differ as to whether or not they employ a unified strategy for coding focus. Some languages use a unified, category-neutral strategy (cf. examples (14) and (15)). Others have category-dependent focus-strategies (cf. example (16)).

Hausa and Hdi are representatives of the first type. These languages have a unified strategy based on the movement strategy for nominal focus (see (7) and (8) above). V- and VP-focus are marked by assimilation to the nominal strategy. In Hausa, focused verbs have to be nominalized before being fronted (Newman 2000). (14a) is a neutral sentence. In (14b), the VP is nominalized (indicated by lengthening of the final vowel) and moved to the sentence initial position (DEP = dependent = a specific auxiliary form obligatory with A'-movement in Hausa).

(14) a.   su-n   bàzamà   *(neutral)*
          3pl-PERF   bolt.away
          'They bolted away.'

     b.   *bàzamàa*   su-kà   yi   *(VP-focus)*
          bolting.away   3pl-PERF.DEP   do
          'They BOLTED AWAY.' (lit. 'Bolting away, they did.')

Hdi inserts a cognate object that is fronted when the verb is in focus (Frajzyngier 2002), cf. (15b) (D:SO = distal extension, point of view of source).

(15) a.   mbàzá-ùgh-mbàzá Pghinta tá   mbàzá   *(neutral)*
          wash-D:SO-wash   P.   OBJ   wash
          'Phinta washed.'

     b.   *mbàzá* mbàzá-ùgh-mbàzá Pghinta   *(V(P)-focus)*
          wash   wash-D:SO-wash   P.
          'Phinta WASHED.' (lit. 'Wash, Phinta washed.')

The second group of languages uses category-dependent focus strategies. In Mupun and Tangale, for instance, focus on nominal expressions is expressed differently from focus on verbs and VPs. In Mupun, focused nominals carry a focus marker 'a' (see (11)), whereas focused verbs reduplicate in addition (Frajzyngier 1993):

(16)   mo cet a   cet lua ne ba mo sur(a)sur kas   *(V(P)-focus)*
       3pl boil FOC boil meat the NEG 3PL fry FOC fry NEG
       'They BOILED the meat, they didn't FRY it.'

As we will show in section 5, in Tangale, at least some focused nominals move to a postverbal focus position (see (8b) above), whereas focused verbs (and VPs) show no sign of movement. Again, there seem to be at least two strategies for focusing a constituent.

The data discussed in this section lead us to conclude that some Chadic languages use different strategies for focusing different syntactic categories. This forces us to refute the Preliminary Hypothesis assumed in (13). Some Chadic languages differ from stress languages in that more than one factor has to be considered in focus marking. In the next section we will analyse the Tangale focus system in detail. The discussion will provide more evidence for the claim that the standard focus theories do not extend directly to all Chadic languages.

5   Predicate Focus in Tangale

In this section, we take a closer look at predicate focus, i.e. V- or VP-focus in Tangale, a Western Chadic language from the Bole-Tangale subbranch.[1] We present the main empirical findings in 5.2. For a better understanding of the following discussion, however, it is necessary to first take another look at (argument) DP-focus in Tangale.

5.1 Existing accounts of focus in tangale

The—to the best of our knowledge—two existing accounts of focus in Tangale (Kenstowicz 1985, Tuller 1992) assume focus to be realised syntactically: The focused DP is moved (sometimes vacuously) to a postverbal position. The two accounts differ only as to the direction of movement.

In Kenstowicz (1985:86), focused (DP-) constituents move to the right and adjoin to S (or S'). In the neutral, all new sentence (17a), the subject is in its

---

[1]   For a general introduction into the grammatical system of Tangale, see Jungraithmayr (1956), as well as the two grammatical sketches in Jungraithmayr (1991) and Kidda (1993).

unmarked sentence-initial position and precedes the verb. When focused, however, the subject moves to a postverbal position (17b).[2]

(17) a.   [$_S$ Malay [$_{VP}$ múdúd-gó]]   *(neutral)*
          M.   die-PERF
          'Malay died.'

     b.   [$_S$ t_1 múdúd-gó] nóŋ$_1$ *(SUBJ-focus)*
                 die-PERF   who
          'Who died?'

In a parallel fashion, direct objects are assumed to move vacuously for reasons that have to do with the different phonological realisation of the perfective aspect marker as *-ug* or *-go* in (18ab):

(18) a.   [$_S$ Kay [$_{VP}$ dob-*ug*   Málay]]   *(neutral)*
          K.   call-PERF M.
          'Kay called Malay.'

     b.   [$_S$ Kay [$_{VP}$ dob-*gó* t_1] nóŋ$_1$]   *(OBJ-focus)*
          K.   call-PERF
          'Who did Kay call?'

While focused (DP-) constituents also move in Tuller's (1992) analysis, the direction of movement is to the left and the focused material left-adjoins to the VP-projection. Since the perfective verb has to move to the inflectional head I[0] for independent reasons, focused constituents nevertheless surface in a postverbal position, as shown for a focused object in (19) (cf. Kenstowicz's 18b).[3]

---

[2]   We abstract away from the open/closed distinction in vowel quality.
[3]   Tuller does not discuss the precise structure of clauses with focused subjects.

(19)   [$_S$ [$_{IP}$ Kay dob-gó [$_{VP}$ nóŋ₁ [$_{VP}$ t$_v$ t₁]]]]   *(OBJ-focus)*
                  K.   call-PERF   who

As indicated above, there is only indirect, namely phonological evidence for the assumption of vacuous movement in the case of focused objects (be it to the left or to the right). The evidence comes in form of a prosodic barrier between V and the focused OBJ that blocks two phonological processes, namely *vowel elision* (henceforth: VE) and *left line delinking*.[4]

In (18b) with a focused object, the prosodic barrier preceding the object blocks VE and the perfective marker must be realised as *-go*. Had VE applied, the perfective marker would have been realised as *-ug*. Kenstowicz (1985:80) defines VE as follows (where ']' marks the end of the stem or word):

(20)   *Vowel Elision (VE)* deletes the final vowel of a stem or a word when in
        close syntactic connection with some following phonological material
        denoted by the X: V $\rightarrow$ $\varnothing$ / _] X

The relevant restriction here is that VE between two elements is possible only when the two elements stand in a close syntactic relation, e.g. head-complement. Application of VE to perfective verbs elides the final vowel of the perfective marker *-go* (cf. 21b). Since the result of elision does not comply with Tangale syllable structure, an epenthetic vowel *-u-* is inserted in a last step (cf. 21c).

(21)   a. mad-*gó* 'read-perf'   >>   b. mad-*g* (after VE)   >>   c. mad-*ug*

---

[4] Kidda (1993:110) speaks of a *strong boundary* in this connection. Apart from vowel
     elision and left line delinking, Kidda (1993:135) cites three more phonological processes,
     namely right line delinking II, decontouring, and P-lowering, which are also blocked at a
     strong boundary before a focused object.

The (non-) application of VE is relevant for the present discussion because it gives us a reliable diagnostic for OBJ-focus. The empirical generalisation is that whenever the object is focused, VE is blocked: OBJ$_{FOC}$ ⟺ *VE. For illustration, VE can apply in the neutral sentence (22a), deleting the final -*o* of the perfective marker. In contrast, VE is blocked with the focused (wh-) objects in (22b) such that the perfective marker surfaces as -*gó*:

(22) a. Áudu mad-*ug*   littáfi.
         A.   read-PERF book
         'Audu read a book.'

   b.   Q: Áudu mad-*gó*/*mad-*ug* náŋ?   A: Áudu mad-*gó*/*mad-*ug* líttáfi.
        A.   read-PERF   what   A.   read-PERF   book
        'What did Audu read?'   'Audu read A BOOK.'

Given the definition of VE in (20), the non-application of VE in (22b) implies that verb and object do not stand in a close syntactic relation when the object is focused. From this Kenstowicz and Tuller conclude that the object must have moved (vacuously) away from the verb.

As mentioned above, the presence of a prosodic barrier before focused objects is also indicated by the blocking of a second phonological process: Left line delinking (henceforth LLD), which has the same domain of application as VE in the postverbal domain (Kenstowicz 1985:82), separates tones that have spread to the right from their original tone-bearing unit (Kenstowicz 1985, Kidda 1993). The effect of LLD is visible in (18a), where the underlying H tone of the perfective marker -*gó* (cf. 23a) has spread onto the object (cf. 23b) before being detached from its original tone-bearing unit by LLD (cf. 23c):

(23)   a. dob-gó Malay   →   b. dob-gó   Málay   →   c. dob-ug   Málay
                  H   H   H   by LLD   H

In (18b), where the object is focused, LLD cannot apply. As a result (and since VE is also blocked before focused objects), the resulting surface form is *dob-gó Málay*, with the perfective marker still being attached to its underlying H-tone. As with VE, the blocking of LLD before a focused object therefore shows the presence of a prosodic barrier before a focused object. Using the same argumentation as with VE, Kenstowicz and Tuller take this prosodic barrier to indicate vacuous movement of the focused object.[5]

Neither Kenstowicz nor Tuller discusses instances of V- or VP-focus, to which we turn in the next section. There, it will emerge that the insertion of a prosodic boundary that blocks VE and LLD plays a more general role in Tangale focus marking than so far assumed.

5.2 Verb (phrase)-focus in Tangale

In this section, we show that predicate focus on the verb or on the entire VP in Tangale is in some cases marked differently from argument DP-focus. Unlike SUBJ-focus, predicate focus in Tangale does not involve movement to a postverbal position. Instead, it is sometimes indicated morphologically by means

---

[5]   By and large, similar facts seem to obtain for subject and object focus in Bole, a closely related SVO-language (see Schuh 2004). Focused subjects appear ex situ (ia), while focused objects (and other focused constituents) remain in situ (ib).

(i)   a.   kàppū   mòrɗo lò?   kàppū   mòrɗo Bamoi
           planted millet who   planted millet B.
           'Who planted millet?' 'BAMOI planted millet.'

      b.   ita à   kòna   làawò lò?   ita à   kòna   làawò Bamoi
           she aux take(fut) child who   she aux take(fut) child B.
           'Whose child will she take?'   'She will take BAMOI'S child.'

As in Tangale, the focus status of objects is indicated by the blocking of a phonological process, namely *low tone raising* (LTR), see Schuh (2004) for details.

of a verbal suffix (5.2.1), or prosodically by the insertion of a prosodic boundary (5.2.2). Thus, there seem to be at least three strategies of focus marking in Tangale: syntactic movement, suffixation, and prosodic phrasing. In addition, we show that V-, VP- and OBJ-focus are often realised identically to the exclusion of SUBJ-focus, arguing against Kenstowicz's (1985) and Tuller's (1992) analyses of OBJ-focus as involving vacuous movement.

In eliciting the various focus markings in Tangale, we used contexts invoking different pragmatic foci (as defined in section 2.1), namely corrective, selective, and new-information focus. The elicited data do not seem to show variation across these contexts, suggesting that focus marking in Tangale (as in stress languages) is insensitive to such pragmatic distinctions.

5.2.1 Morphological focus marking

With some intransitive verbs, V(P)-focus is marked morphologically by means of a verbal suffix -*i*.[6] This is shown in (24b), where the verb (or the entire VP) is in focus and the suffix is added after the perfective suffix -*go*. In contrast, no special focus-suffix is added in neutral, all new contexts (24a):

(24) a. Fátíma wur-go.   *(neutral)*
         F.   laugh-PERF
         'Fatima laughed.'

     b. Q: Mairo yaa-gó náŋ? A: Mbáastám   wur-gó-*i*.   *(V(P)-focus)*
         M.   do-PERF what   she   laugh-PERF-FOC
         'What did Mairo do?'   'She LAUGHED.'

---

[6] For reasons unclear to us, this focus marking device does not seem to occur with all intransitive verbs. Also, *i*-suffixation exhibits a certain degree of optionality even with those verbs on which it can occur in principle.

This is a focus strategy differing from the one observed for focused subjects, which involved movement to a postverbal position, as shown in (17b). Unlike in stress languages, there are thus at least two focus strategies in Tangale, one of them (suffixation) seemingly reserved for intransitive verbal predicates.

5.2.2 Prosodic focus marking

Prosodic focus marking is used with transitive verbs or VPs. It turns out that the phonological processes of vowel elision (VE) and left line delinking (LLD) on perfective verbs are blocked not only with focused objects (see section 5.1), but also with focused verbs or VPs. (25a) is an already familiar example with OBJ-focus. The crucial cases are (25b), with VP-focus, and (25c), with V-focus.

(25) a. Q: What did Laku sell?   *(OBJ-focus)*
         A: Lak wai-gó   lánda
             L.   sell-PERF dress
             'Laku sold [A DRESS]~FOC~.'

     b. Q: What did Laku do?   *(VP-focus)*
         A: Lak waig-ó   lánda
             L.   sell-PERF dress
             'Laku [sold A DRESS]~FOC~.'

     c. Q: What did Laku do at the market?   *(V-focus)*
             Did she buy a dress or did she sell a dress?
         A: Lak wai-gó   lánda
             L.   sell-PERF dress
             'Laku [SOLD]~FOC~ a dress.'

In all three cases, the perfective verb appears in its non-elided form *wai-gó*, and the H-tone has not been detached from the perfective marker *-gó* by LLD. The blocking of both VE and LLD indicates the presence of a prosodic phrase boundary after the verb, which makes the three cases identical in syntactic and

phonological structure. In section 5.3, we will show on the base of exemplary pitch tracks that the three foci in (25a-c) do not appear to be distinguished by other prosodic means (prosodic breaks, tone raising, etc.) either.

The prosodic phrase boundary after the verb in (25b) cannot be the direct result of moving the VP as a whole, since the boundary is inside the VP. Nor can the prosodic phrase boundary in (25c) be the result of verb movement for principled reasons. Obviously, the verb in (25c) has not moved to the right, adjoining to S (see Kenstowicz 1985). What about movement to the left, say to the head of a functional projection FocP? According to Tuller (1992), perfective verbs must, focused or not, move to the inflectional head I^0 in order to support the perfective suffix. Tuller (1992:317) further assumes that verb traces in Tangale are unable to assign case to their direct object. Therefore, whenever the verb moves, the object has to move along with it (presumably after incorporating into the verb) for reasons of case. Hence, if the verb moved to Foc^0 on its way to I^0 in (25c), the object would move along, preserving the close syntactic relation between the two elements (recall that VE only applies between locally related elements). As a result, VE should not be blocked in (25c).

The alternative assumption that the verb moves to I^0 on its own, leaving its object behind in its base position, makes wrong predictions as well. After V-(to-Foc-)to-I movement, verb and object would no longer stand in a close syntactic relation such that VE should be blocked. However, since movement to I^0 is assumed to take place whether or not the verb is in focus, we would expect VE to be blocked in all perfective sentences. This prediction is falsified by (26), from Kidda (1993:122), where VE applies in a neutral all new sentence:

(26)   Lak šwad-ùg   yiláà
        L.   hit-PERF   Y.
        'Laku hit Yila.'

We conclude that the insertion of a prosodic phrase boundary is a focus marking device independent of movement. Focus on the VP in (25b) and on the verb in (25c) are marked by inserting a phrase boundary at PF. No previous syntactic movement is necessary. But given this, we no longer have to assume that the prosodic phrase boundary showing up with OBJ-focus in (25a) is the result of vacuous movement, as argued by Kenstowicz (1985) and Tuller (1992) (see section 5.1). Rather, V-focus, VP-focus and OBJ-focus seem to be marked by the same formal device, namely by inserting a prosodic phrase boundary to the right of the verb. This phrase boundary signals that some element of the VP, or the entire VP is in focus. Tangale thus differs from stress languages, in which narrow V-focus is marked differently from narrow OBJ-focus by stress placement on the verb or the object, respectively.

In contrast, SUBJ-focus with transitive verbs is again marked by syntactic movement. As in the intransitive sentence (17b), the focused subject in (27) has moved from its default preverbal position to a postverbal position.

(27) t_1 way-ug   land-í   nóŋ$_1$?   *(SUBJ-focus)*
           sell-PERF dress-the who
       'Who sold the dress?'

Summing up, there seem to be at least three focus strategies in Tangale, namely syntactic movement, *i*-suffixation, and prosodic phrasing. These strategies are in part dependent on the syntactic category or the grammatical function of the focused constituent. Syntactic movement seems to be reserved for focused subjects, while *i*-suffixation is reserved for (intransitive) verbal predicates. With transitive verbs, instances of V-, VP- and OBJ-focus are not formally distinguished, leading to focus ambiguity.

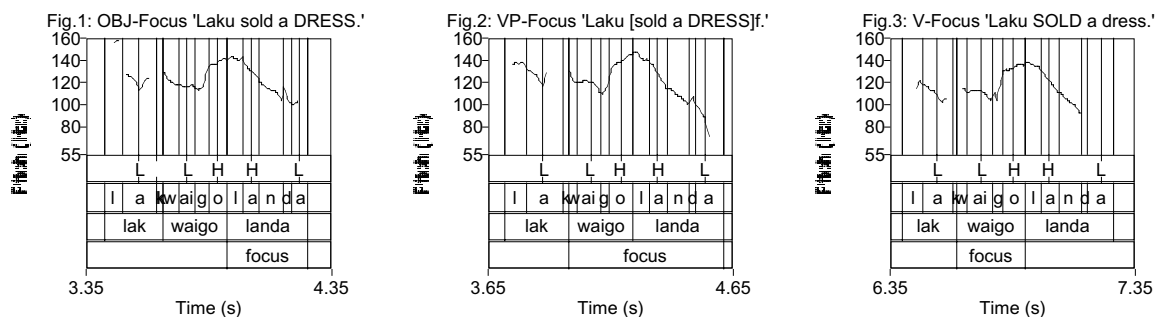### 5.3 Focus and prosody in the perfective aspect

In the previous section, we showed that a prosodic boundary is inserted after the verb not only with focused objects, but also when a verb or a VP is focused. The existence of this prosodic boundary is witnessed by the fact that the two phonological processes of vowel elision (VE) and left line delinking (LLD) are blocked. This raises the question if there are any other prosodic clues, such as intonational breaks, boundary tones, tone raising, register height etc., which would formally distinguish the three different focus structures.

In order to establish if there are any significant prosodic differences between structures with VP-, V-, or OBJ-focus, we conducted a production experiment. We compiled a list of in total 170 Tangale sentence pairs with different focus structures (VP-, V-, OBJ-, and all-new focus) in three different aspects (perfective, progressive, future).[7] The individual pairs consisted of a trigger sentence and a target sentence. In most cases, the trigger sentence was a question that determined the focus structure of the corresponding answer, the target sentence. For instance, the question *Lak yaa-go nang?* 'L. do-PERF what = What did Laku do?' determines that the answer will contain a VP-focus. The 170 sentences were randomly mixed with regard to focus structure and aspect in order to prevent repetitive effects. The consultant was then asked to read each sentence pair aloud. The recording was converted into a WAV.-file, which was then analysed with PRAAT. For each target sentence, we extracted the F0 tracing in order to check for differences in intonation.

---

[7]   42 sentence pairs tested V-, VP-, and OBJ-focus in the perfective aspect. 49 sentence pairs tested V-, VP, OBJ, and all-new focus in the progressive aspect. 49 sentence pairs tested V-, VP, OBJ, and all-new focus in the future (=long progressive) aspect. 15 sentences tested the association of the focus particle *nǘm* 'only' with V-, VP-, and OBJ-focus in the perfective aspect. An additional 15 sentences tested the association of the focus particle *nǘm* 'only' with V-, VP-, and OBJ-focus in the progressive aspect.

Looking at the phonetic realisation of the 42 perfective sentences that were recorded, we could find no significant prosodic differences between V-, VP-, and OBJ-focus.[8] The three pitch contours for (25a-c) are given in figure 1-3.


Fig.1: OBJ-Focus 'Laku sold a DRESS.'   Fig.2: VP-Focus 'Laku [sold a DRESS]f.'   Fig.3: V-Focus 'Laku SOLD a dress.'

The three tone contours appear to be virtually identical. In all three structures, H-tone has spread from the perfective marker *-gó* onto the first syllable of the object. In all three structures, the H-tone has not been detached from its original tone-bearing unit, the perfective marker *-gó*. The three low tones are either lexical tones (*lak*), or derived by the general tone rules *m(orphological)-lowering* (lowers the tone of the verb before the suffix *-gó*) and *p(honological)-lowering* (lowers the second tone of the object before a pause, presumably due to a low boundary tone L% at the edge of the intonational phrase), see Kidda (1993) for discussion. In addition, there is no evidence for most of the intonational processes that tone languages commonly use in order to indicate structural (here: information structural) differences (see Yip 2002:260). The entire pitch register and the pitch range of the three utterances are the same.

---

[8]   In the analysis, we have only looked for differences at the *phonetic surface* that would help to distinguish the different foci. We do not exclude the possibility that there could be phonological differences underlyingly, which - for some reason - are neutralized at the phonetic surface (see the remarks below fig. 4-6, which go in the same direction). However, it is not clear to us why focus marking in a language should be organised in such a way that its results are never, or hardly ever perceivable. Also bear in mind that a purely qualitative analysis such as presented here may miss certain significant differences, and should be supplemented by a quantitative analysis and a perception study.

Also, there is no sign of additional boundary tones inserted at the edge(s) of the respective focus domains. Finally, there are no intonational breaks either before or after the focus domain, nor are there any differences in vowel length.

The only discernible difference in Fig. 1-3 concerns the relative height of the two adjacent H-tones. In the case of VP-focus (fig.2), the second H-tone on *lán* seems to be lower than the first H-tone on *-gó*, whereas it seems to be slightly higher in the case of OBJ-focus (fig.1) and V-focus (fig.3). One could therefore speculate whether the lower second H-tone in the case VP-focus is not the result of *downdrift/downstep* or *declination* (Yip 2002:262), which in this case would not be blocked by an intervening focus boundary.[9] In the case of OBJ-focus and V-focus, downdrift/downstep or declination would be blocked by the intervening focus boundary, resulting in a reset of the next H tone to the original level. Apart from the fact that the realisation of V-focus and OBJ-focus would still be identical (unlike in stress languages), such an hypothesis is not supported by additional data.

Fig. 4-6 show that the prosodic realisation of the three different foci in the sentence *Lak saa-gó foo* 'L. eat-PERF mush = Laku ate mush' does not differ.

Fig.4: OBJ-Focus 'Laku ate MUSH.'   Fig.5: VP-focus 'Laku [ate MUSH]f.'   Fig.6: V-focus 'Laku ATE mush.'

---
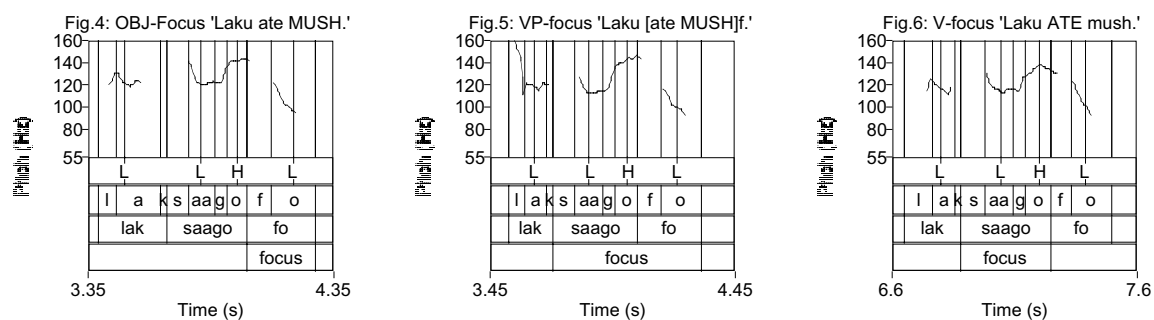
[9]   There is no discussion of downdrift/downstep or declination in Tangale in Kidda (1993). It appears likely, though, that some such process is active in Tangale, as it is in Hausa. Inkelas & Leben (1990) show that downstep in Hausa can be interrupted by smaller phonological phrase boundaries within the intonational phrase, such that the next H tone is raised.

Because the monosyllabic object *foo* occurs before a pause, p-lowering will lower its tone independent of other tonal processes (H-tone spread) that may have applied before, thereby neutralizing any potential differences in tone height. As a result, the tone of the object will always be lower than that of the perfective marker *-gó* (notice again that LLD has not applied to *-gó*) and the F0 tracings of the various focus structures are identical.

Finally, fig. 7-9 show that the same holds for the sentence *Lak bal-gó wásíika* 'L. write-PERF letter = Laku wrote a letter' with a trisyllabic object, where potential differences in tone height are not neutralized by final p-lowering.

Fig.7: OBJ-Focus 'Laku wrote a LETTER.' Fig.8: VP-Focus 'Laku [wrote a LETTER]f.' Fig.9: V-Focus 'Laku WROTE a letter.'
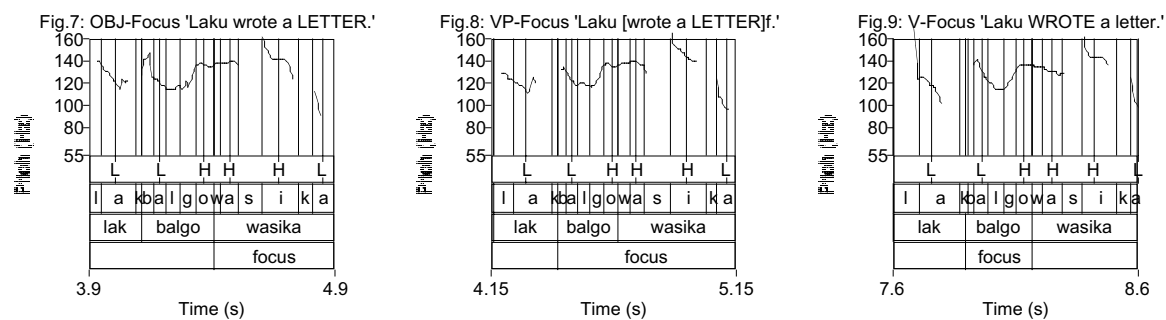
In the absence of further evidence, we therefore conclude that prosody is not used in order to disambiguate V-, VP-, and OBJ-focus in perfective sentences in Tangale.[10] The same will be shown for the progressive aspect in section 6.2.

---

[10] There may be a potential methodological problem lurking here, which has to do with the general set up of the production experiment. Questions triggering VP-focus in the answer are of the same general form than questions triggering OBJ-focus, namely of the form *x yaa-go nang?* 'X do-perf what = what did x do?'. Strictly speaking, only the object is focused in such a question. In principle, it is possible that a requirement on *phonological parallelism* between (focused) answers and the trigger questions is operative in Tangale. If so, we would expect no phonological differences between sentences with OBJ-focus and sentences with VP-focus despite their differences in focus structure. This brings out nicely the general methodological problem of using linguistic triggers in eliciting linguistic data.

6 An Alternative Solution: Subjects vs Non-Subjects?

6.1 Focus theories revisited

In section 2, we have seen that focus in stress languages can be captured by a fairly simple model that considers only one factor, namely stress.

(28)   Focus model for stress languages (based on Selkirk 1995):
        CONSTITUENT STRESSED → focus/new, otherwise old information

In sections 4.1 and 4.2, it was then shown that this mono-factorial model of the standard analysis can be extended to some Chadic languages, such as Hdi. In Hdi, focus marking of all categories is assimilated to the nominal strategy, such that only movement has to be considered:[11]

(29)   Focus model for Hdi:
        CONSTITUENT MOVED → focus/new, otherwise old information

Due to the lack of information on predicate focus in most Chadic languages, it remains to be seen if a mono-factorial analysis can be extended to those languages that employ only one strategy for marking nominal focus (see 3.1).

Given the discussion in section 5, it is clear that focus marking in Tangale is more complicated, and cannot easily be captured by mono-factorial models like those sketched in (28) and (29). (The same may hold for Pero, which also

---

[11] In Hausa, another language that assimilates marking of predicate focus to the nominal strategy of focus movement, the situation is complicated by the fact that it also allows for in situ focus (see the remarks in section 4.1). If so, checking of whether or not a constituent has moved to the initial position is insufficient for determining the precise information structural status of a constituent as being old information: An element could still be in focus (in situ) without having moved. Interestingly, in situ focus in Hausa displays a subject-object asymmetry similar to that observed for Tangale in the main text. Unlike objects, subjects cannot be focused in situ (see Green & Jaggar 2002).

makes use of more than one focus strategy, namely movement and prosodic phrasing, see section 4.1). Based on the data in 5.1 and 5.2, a model of focus marking in Tangale would have to consider at least three factors as shown in (30):

(30)   Focus model for Tangale:
       if CONSTITUENT MOVED   → SUBJ-focus, otherwise
            if i-SUFFIXATION   → intransitive V(P)-focus, otherwise
                 if PROSODIC BOUNDARY → V, VP-, OBJ-focus, otherwise
                      old information or neutral

It seems, then, that focus marking in Tangale is a complex process that requires a more complex theory of focus.

6.2 Focus in the progressive aspect

The picture of focus marking in Tangale gets even more complicated when verbal aspects others than the perfective are considered. In the progressive, there are no discernible differences at all between neutral, i.e. all-new sentences on the one hand (31), and sentences with OBJ-focus, or VP-focus, or V-focus, on the other (32a-c). In all cases, VE obligatorily deletes the final vowel on the verbal noun *balli > ball*.[12]

(31)   Lakú n   ball   wasíika   *(neutral)*
       L.   PROG writing letter
       'Laku is writing a letter.'

(32) a. Q: Lakú n   ball   náŋ? A: Lakú n   ball   wasíika *(OBJ-focus)*
        L.   PROG writing what   L.   PROG writing letter
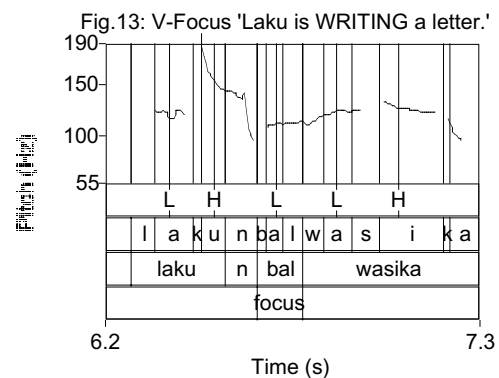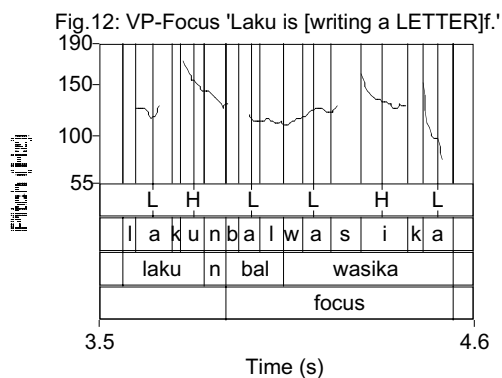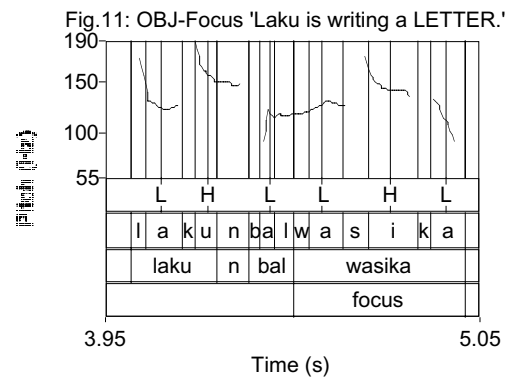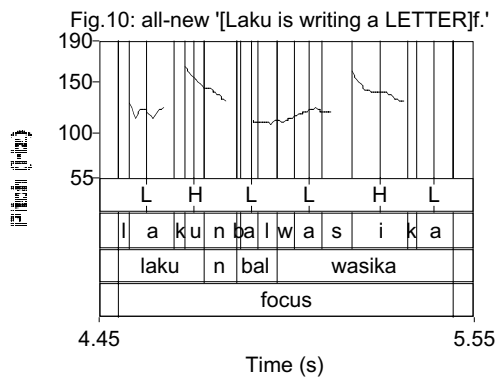   'What is Laku writing?'   'Laku is writing A LETTER.'

---

[12] Here, our elicited data are not in accordance with Kidda's claim (1993:127) that VE in the progressive is blocked before focused objects, as it is in the perfective.

   b. Q:Lakú n   yaaj   náŋ? A: Lakú n   ball   wasíika *(VP-focus)*
       L.   PROG doing what   L.   PROG writing letter
       'What is Laku doing?'   'Laku is [writing A LETTER]_F.'

   c. Q:Lakú n   ball   wasíika yá mad   wasíika?
       L.   PROG writing letter   or reading letter
       'Is Laku WRITING a letter or READING a letter?'

    A:Lakú n   ball   wasíika   *(V-focus)*
       L.   PROG writing letter
       'Laku is WRITING a letter.'

The reason for this formal identity has to do with the fact that the focus marking device for OBJ-focus and V(P)-focus in Tangale, i.e. the insertion of a prosodic phrase boundary between verb and object (see 4.2), is bled by the syntactic structure of the progressive plus the general conditions on VE. As in Hausa, verbs are nominalised and form an N-N-complex with their direct object in the progressive aspect. Kenstowicz (1985) shows that VE obligatorily applies in such N-N-configurations, presumably because the two N-elements stand in a close syntactic relation. But if VE must apply obligatorily, it can no longer serve as a diagnostic for OBJ-focus and V(P)-focus in the progressive aspect. In other words, narrow focus on V(P) or object does not seem to be explicitly marked at all in the progressive, resulting in an underspecification of focus.[13]

     Again, this conclusion is supported by a closer inspection of the pitch contours associated with the different focus structures in (31) and (32a-c). As shown in the following figures, the pitch contours of neutral focus (fig.10), OBJ-focus (fig. 11), VP-focus (fig.12) and V-focus (fig.13) appear to be identical in all relevant aspects.

---

[13] The same holds for the future, or long progressive, which is identical in syntactic structure.

Fig.10: all-new '[Laku is writing a LETTER]f.'

	L	H	L		L		H	L
l	a	k̠u	n	b̠a	l	w	a	s	i	k̠	a

laku | n | bal | wasika

focus

4.45 — 5.55

Time (s)

Fig.11: OBJ-Focus 'Laku is writing a LETTER.'

	L	H	L		L		H	L
l	a	k̠u	n	b̠a	l	w	a	s	i	k̠	a

laku | n | bal | wasika

focus

3.95 — 5.05

Time (s)

Fig.12: VP-Focus 'Laku is [writing a LETTER]f.'

	L	H	L		L		H	L
l	a	k̠u	n	b̠a	l	w	a	s	i	k̠	a

laku | n | bal | wasika

focus

3.5 — 4.6

Time (s)

Fig.13: V-Focus 'Laku is WRITING a letter.'

	L	H	L		L		H
l	a	k̠u	n	b̠a	l	w	a	s	i	k̠	a

laku | n | bal | wasika

focus

6.2 — 7.3

Time (s)

It seems, then, that focus marking in Tangale is not only a complicated process, but also an underspecifying process with systematic gaps. In certain aspects, narrow focus (be it on the OBJ, on VP, or on V) does not seem to be indicated at all. This is a surprising result given that the theories of focus generally assume that (narrow) focus must be marked somewhere on the focused constituent.

Interestingly, the only constituent in Tangale that can unambiguously be marked for focus even in the progressive and future aspect is the subject. As in (17b) and (27) above, the subject occurs again in a postverbal position.[14]

---

[14] When the subject is focused, the word order (nominalised) V >> OBJ >> SUBJ is often changed by making the object the (optional) sentence-initial topic of the utterance. In such a case, a pronominal suffix -i is added to the nominalised verb, as illustrated in the answer in (33). It remains to be seen if there exists more than an accidental homophonic relationship between the neutral pronominal suffix -i and the focus marker -i discussed in section 4.2.1.

(33)   Q: bal   wasíika-i nóŋ?   A: (wasíika-i) ball-í   Músa
       writing letter-the who   letter-the   writing-it M.
       'Who is writing the letter?'   'MUSA is writing the letter.'

The data in (31)-(33) give rise to the following empirical generalisation:

(34)   In Tangale, focus marking is fully grammaticalised only on subjects. On
       all other constituents, focus is only sporadically marked and relies
       heavily on pragmatic resolution.

The generalisation in (34) is a more drastic version of the hypothesis that focus
on different syntactic categories is marked differently, which was argued for in
sections 4 and 5. On some syntactic categories, focus may be left unmarked. If
correct, the generalisation in (34) allows for a significant simplification in the
focus marking system of Tangale, as sketched in (35).

(35)   Alternative focus model for Tangale:
       CONSTITUENT MOVED → SUBJ-focus, otherwise the interpretation of
                           elements as focused or not is pragmatically resolved.

The underspecification-based model in (35) seems to be all that can be said
about Tangale focus marking in the progressive and future aspect, and perhaps
even in general.[15] Interestingly, there is additional evidence in favour of (35).
This evidence comes from the behaviour of the focus particle *núm* 'only', to
which we turn now.

---

[15] If (35) is an adequate model of focus marking in Tangale in general, the question arises
why focus can or should ever be marked on constituents other than the subject, as was
shown in sections 5.1 and 5.2. At the moment, we have no conclusive answer to this.

6.3 Association with focus

The hypothesis that focus marking in Tangale does not differentiate between V-, VP, and OBJ-focus in most cases is supported by the behaviour of the focus particle *núm* 'only'. Semantically, *núm* can associate either with a focused object (36a), or with a focused VP (36b), or with a focused verb (36c). Syntactically, however, it can only combine with nominal (DP) expressions like its Hausa counterpart *sái* and unlike its English counterpart *only*. For this reason, the different narrow foci in (36a-c) come with identical syntactic structures.

(36) a. N fad-go   *núm* littáfi-i,   n fad-ug   wam gáayi-m   *(OBJ-focus)*
        I buy-PERF   only book-the I buy-PERF   s.th.else-NEG
        'I bought only THE BOOK, I bought nothing else.'

   b. N fad-go   *núm* littáfi-i,   n yaa-g   wam gáayi-m   *(VP-focus)*
        I buy-PERF   only book-the I do-PERF   s.th.else-NEG
        'I only bought THE BOOK, I did nothing else.'

   c. N fad-go   *núm* littáfi-i,   fon di n mad-go-m   *(V-focus)*
        I buy-PERF   only book-the but yet I read-PERF-NEG
        'I only BOUGHT the book, but I have not read (it) yet.'

In addition, the pitch tracks for (36a-c) in fig. 14-16 suggest, once again, that there are no prosodic differences either. In each case, presence of the focus particle *núm* effects a rise from the preceding H-tone on *-gó* to an extra high tone on *núm*. It also leads to a considerable raise in the pitch register of the utterance. Otherwise, *núm* appears to be tonally 'opaque' in that it does not spread its H-tone onto the next tone bearing unit *li*.

Fig.14: OBJ-Focus 'I bought only the BOOK...'   Fig.15: VP-Focus 'I only bought the BOOK...'   Fig.15: V-Focus 'I only BOUGHT the book...'
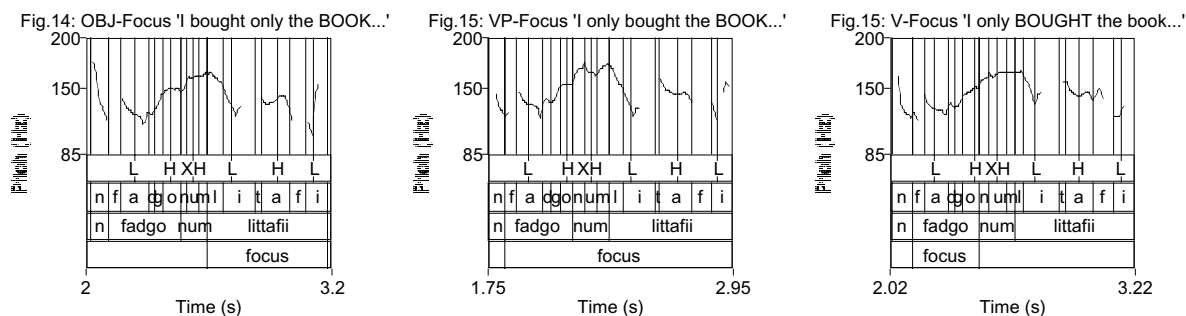
Setting aside the tonal properties of *núm*, we conclude that the presence of a focus-sensitive particle such as *núm* does not help to distinguish OBJ-, VP-, or V-focus, neither syntactically nor phonologically.[16] The sentences in (36a-c) with *núm* are as ambiguous with respect to focus structure as are their counterparts without (see section 6.3).

In contrast, the focus particle *núm* can only combine and associate with focused subjects when these have moved to postverbal position.

(37) a.   Landa   pad-go   *núm Laku*
           dress   buy-PERF   only L.
           'Only LAKU bought a dress. (Nobody else bought a dress).'

     b. * *Núm   Laku*   pad-go   landa
           only   L.   buy-PERF   dress

Concluding, the data from association with focus with the focus-sensitive particle *núm* support the hypothesis that there is a fundamental asymmetry between focus marking of subjects and focus marking of non-subjects. Only association with a focused subject is marked unambiguously (by displacing the

---

[16] Association with focus with *núm* has other interesting characteristics with theoretical repercussions. Due to the fact that *núm* can only combine with nominal (DP-) expressions, association with focus does not seem to be subject to c-command in Tangale, and possibly Chadic languages in general. This means that the c-command requirement for association with focus (Büring and Hartmann 2001) cannot be a language universal. Possibly, the

focused element). Association with other focused constituents (OBJ, VP, V) is marked ambiguously and left open for pragmatic resolution. Altogether, our findings support the claim that focus marking may not be fully grammaticalised in Tangale.

6.4 Cross-linguistic parallels

The model in (35) draws a sharp line between subjects and non-subjects when it comes to focus marking. It singles out focused subjects as being in special need of explicit focus marking. Intuitively, the reason for this apparent subject bias in the Tangale focus system seems clear. The (default) preverbal subject position triggers a topic interpretation (see Givon 1976). Therefore, if a subject is to be interpreted as focus (and not as topic) something special has to be done. In the Tangale case, the subject has to be dislocated.

A comparable special status for focused (wh-) subjects has been observed for a number of languages both within and outside the Chadic language family. For instance, in the Bantu languages Kinyarwanda, Dzamba, and Kitharaka, and also the Austronesian languages Malagasy, Tagalog, and Javanese, wh-subjects have to move, whereas wh-objects can remain in situ (see Sabel & Zeller, to appear, and references therein). Looking again at the Chadic languages, it was mentioned in fn. 11 in section 6.1 that focused objects in Hausa can remain in situ whereas focused subjects have to move (Green & Jaggar 2002). Similarly, focused subjects must move in Bole, whereas focused objects appear to remain in situ (see fn. 5 in section 5.1). Finally, focused subjects require special TAM's (tense-aspect-mood markers) in Miya, whereas focused objects can only be

requirement only holds for languages like stress accent languages, which have the means to grammatically mark individual narrow foci.

identified indirectly by the absence of the totality marker (see (11) above and Schuh (1998) for more discussion).

Hopefully, future work will show more clearly if and to what extent the distinction between subjects and non-subjects plays a central role in the focus systems of the Chadic languages. At any rate, it appears inevitable to us that more attention be paid to the realisation of focus on non-nominal categories.

7 Conclusion

In this paper, we have investigated nominal and verbal focus marking in various Chadic languages, in particular in Tangale. While it seems possible to extend the standard mono-factorial analyses of stress languages to some of the Chadic languages (e.g. to Hdi), the focus systems of other Chadic languages seem to be more complex. Our investigation of the Tangale focus system has shown that three different factors play a role in the perfective aspect. We also showed that narrow foci on object, verb, and VP are not formally distinguished in Tangale. In the progressive aspect, a special focus marking on V, VP, or OBJ appears to be absent altogether, resulting in an underspecification of focus. Given this underspecification, an alternative solution would be to keep the focus system of Tangale simple (assuming only a single distinction between SUBJ- and non-SUBJ-focus) at the cost of shifting the major burden of focus resolution to the pragmatic system.

References

Büring, Daniel. 2004. Focus Projection and Default Prominence. To appear in: Valéria Molnár and Susanne Winkler (eds.) The Architecture of Focus. Studies in Generative Grammar. Berlin, New York: Mouton de Gruyter.

Büring, Daniel and Katharina Hartmann. 2001. The Syntax and Semantics of Focus-Sensitive Particles in German. *Natural Language and Linguistic Theory* 19, 229-281.

Frajzyngier, Zygmunt. 1989. *A Grammar of Pero*. Berlin: Reimer.

Frajzyngier, Zygmunt. 1993. *A Grammar of Mupun*. Berlin: Reimer.

Frajzyngier, Zygmunt. 2001. *A Grammar of Lele*. Stanford: CSLI Publications.

Frajzyngier, Zygmunt. 2002. *A Grammar of Hdi*. Berlin: Mouton de Gruyter.

Givón, Talmy. 1976. Topic, pronoun and grammatical agreement. In: C. Li (ed.) *Subject and Topic*. London/New York: Academic Press, 149-88.

Green, Melanie & Philip J. Jaggar. 2002. Ex-situ and In-situ Focus in Hausa: Syntax, Semantics, and Discourse. To appear in: Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky (eds.) *Research in Afroasiatic Grammar III*. Amsterdam: John Benjamins Publishing Company.

Inkelas, Sharon & Will Leben. 1990. Where phonology and phonetics intersect: the case of Hausa intonation. In J. Kingston & M. Beckman (eds.) *Between the Grammar and Physics of Speech: Papers in Laboratory Phonology I*. Cambridge: Cambridge University Press, 17-34.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jaggar, Philip J. 2001. *Hausa*. Amsterdam: John Benjamins Publishing Company.

Jungraithmayr, Herrmann. 1956. Untersuchungen zur Sprache der Tangale in Nordost-Nigerien. Unpublished Doctoral Dissertation. Hamburg.

Jungraithmayr, Herrmann. 1991. *A Dictionary of the Tangale Language (Kaltungo, Northern Nigeria) with a Grammatical Introduction*. Berlin: Reimer.

Kenstowicz, Michael. 1985. The Phonology and Syntax of WH-Expressions in Tangale. In *Studies in the Linguistic Sciences* 15, 79-91.

Kidda, Mairo E. 1993. *Tangale Phonology. A Descriptive Analysis*. Berlin: Reimer.

Leben, W. R., S. Inkelas and M. Cobler (1989) Phrases and Phrase Tones in Hausa. In: P. Newman and R. D. Botne (eds.) *Current Approaches to African Linguistics*. Dordrecht: Foris, 45-61.

Ma Newman, Roxanna. 1971. *A Case Grammar of Ga'anda*. PhD Dissertation, UCLA, Los Angeles.

Newman, Paul. 2000. *The Hausa Language*. New Haven: Yale University Press.

Rooth, Mats. 1985. *Association With Focus*. PhD Dissertation, UMass, Amherst.

Ruelland, Suzanne. 2000. Topicalisation et Focalisation en Tupuri. In B. Caron (ed.), *Topicalisation et focalisation dans les langues africaines*. Louvain-Paris: Peeters, 135-159.

Sabel, Joachim & Jochen Zeller. To appear. Wh-question formation in Nguni. In *Proceedings of the 35th Annual Conference on African Linguistics*.

Schuh, Russell G. 1998. *A Grammar of Miya*. Berkeley, Los Angeles, London: University of California Press.

Schuh, Russell G. 2004. Draft chapters and sections of a forthcoming Reference Grammar of Bole, downloadable at http://www.humnet.ucla.edu/ humnet/aflang/Bole/ bole_papers.html

Schwarzschild, Roger. 1999. Givenness, AvoidF and Other Constraints on the Placement of Accent. *Natural Language Semantics* 7, 141-177.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 1995. Sentence Prosody: Intonation, Stress, and Phrasing. In J. Goldsmith (ed.) *Handbook of Phonological Theory*. Oxford: Blackwell.

Truckenbrodt, Hubert. 1995. *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. PhD Dissertation, MIT. Published 1999 by MITWPL.

Truckenbrodt, Hubert. 1999. On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry* 30, 219-255.

Tuller, Laurice 1987. Variation in Focus Constructions. In Z. Frajzyngier (ed.) *Current Progress in Chadic Linguistics. Proceedings of the International Symposium on Chadic Linguistics*. Amsterdam: John Benjamins, 9-33.

Tuller, Laurice. 1992. The Syntax of Postverbal Focus Constructions in Chadic. *Natural Language and Linguistic Theory* 10, 303-334.

Xu, Liejiong. 2004. Manifestation of Informational Focus. *Lingua* 114, 277-299.

Xu, Yi. 1999. Effects of Tone and Focus on the Formation and Alignment of f_0 Contours. *Journal of Phonetics* 27, 55-105.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Katharina Hartmann
Institut für deutsche Sprache und Linguistik
Humboldt-Universität zu Berlin
Unter den Linden 6 / Sitz: Schützenstr. 21
10099 Berlin
e-mail: K.Hartmann@rz.hu-berlin.de

Malte Zimmermann
SFB 632, Projekt B2
Humboldt-Universität zu Berlin
Unter den Linden 6 / Sitz: Hausvogteiplatz 5-7
10099 Berlin
e-mail: malte.zimmermann@staff.hu-berlin.de

ANNIS: A Linguistic Database
for Exploring Information Structure

Stefanie Dipper[*], *Michael Götze*[*], *Manfred Stede*[*] *and Tillmann Wegst*[†]

[*]Universität Potsdam and [†]Saarbrücken

In this paper, we discuss the design and implementation of our first version of the database 'ANNIS' ('ANNotation of Information Structure'). For research based on empirical data, ANNIS provides a uniform environment for storing this data together with its linguistic annotations. A central database promotes standardized annotation, which facilitates interpretation and comparison of the data. ANNIS is used through a standard web browser and offers tier-based visualization of data and annotations, as well as search facilities that allow for cross-level and cross-sentential queries. The paper motivates the design of the system, characterizes its user interface, and provides an initial technical evaluation of ANNIS with respect to data size and query processing.

1 Introduction

Information structure (IS) is an area of linguistic investigation that has given rise to a multitude of terminologies and theories that are becoming more and more difficult to survey. The basic problem is that IS-related phenomena often can be observed only indirectly on the linguistic surface and hence invite competing interpretations and analyses that are tailored to the needs and the flavours of the participating researchers. Thus, in contrast to syntax, where different approaches can be—more or less—systematically compared, with IS it is often not even clear whether two theories compete to describe the same phenomenon or are in fact complementary to each other, characterizing linguistic regularities on different levels of description.

In 2003, a long-term research infrastructure ('Sonderforschungsbereich', henceforth 'SFB') has been established at Potsdam University and Humboldt-

University Berlin.[1] Its idea is to investigate the various facets of IS from very different perspectives and to contribute to a broader and more general understanding of the phenomena by bringing the various results together and promoting the active exchange of research hypotheses. Participating projects (see Section 2) provide empirical data analyses that will serve as the basis for formulating theories that aim at advancing the state of the art and overcoming the unpleasant situation characterized above.

An important prerequisite for this long-term and multi-disciplinary approach is the ability to *annotate* the data with appropriate information and to collect the variety of data in a single, uniform database.[2] Given the present situation, annotation sets cannot be presumed to be identical—different researchers will first start out with their own favourite terminology. The convergence of the annotation sets is an important goal for the SFB, and the idea is that this process can be actively promoted by making the interim analyses of the various projects accessible, to invite comparison and possibly revision. Specific working groups dedicated to various levels of analysis are in charge of monitoring this process.

In this paper, we discuss the design and implementation of our first version of the database 'ANNIS' ('ANNotation of Information Structure'). Section 2 provides some more details about the SFB and summarizes the particular requirements that this research scenario places on developing the database. Section 3 explains the architecture, user interface, and query facilities of the current implementation. Then, Section 4 illustrates the operation of ANNIS with an example. Section 5 presents an evaluation of the current state of the database. In Section 6, we compare our approach to related work, and Section 7 discusses our plans for future extensions.

---

[1] `http://www.ling.uni-potsdam.de/sfb/`
[2] For a comparative evaluation of various annotation tools, see Dipper et al. (2004).

2   The SFB

The SFB consists of 13 individual research projects from disciplines such as theoretical linguistics, psycholinguistics, first and second language acquisition, typology, and historical linguistics. Following the overarching objective of providing a clearer picture of information structure, several of the projects are involved in collecting and analyzing empirical data. Here are some examples of such activities.

Semantics and IS   One project examines the relation between *quantifier scope* and IS. Data is annotated with semantic features such as quantifier scope, referent identifiability, and definiteness. Another project investigates interactions between semantic focus evaluation, discourse anaphoricity, and presupposition.

IS and discourse structure   One project is interested in the effects that rhetorical relations and discourse structure in general can have on the prosodic structure of spoken discourse. The data to be annotated with corresponding features are radio news broadcasts.

Focus in African languages   Two projects examine the phenomenon of focus in different Western African languages. Both carry out field studies for collecting data, which is later being annotated.

Diachronic change   One project investigates the evolution of the verb-second phenomenon, which occurred in certain Germanic languages only (e.g., it did in Modern German, but not in Modern English). Based on manuscripts of Old High German and Old English, the role of IS in this evolution will be studied.

Typology of information structure   One project seeks to develop a typology of the means for expressing IS. In close cooperation with the other projects, a

questionnaire is being developed that will serve as a basis to collect language data relevant for IS from speakers of typologically diverse languages.

2.1 The data

As pointed out above, the individual projects apply different means in collecting data, and they focus on different aspects of IS. Hence, both the *primary* data (i.e., the language data that is collected) as well as the *secondary* data (i.e., the annotations to the primary data) of these projects differ in several respects.

Primary data The source data can consist of recorded speech, or videos of spoken monologues or dialogues. Furthermore, some projects work with written texts, either in digital form or as original manuscripts. A special case is the above-mentioned *questionnaire*, whose primary data are answers to questions. Generally, the data is taken from diverse languages, many of which do not make use of the Latin character set.

Secondary data Languages differ with respect to the means they exploit to express IS (e.g. stress, word order, particles). Depending on the objectives of the individual project, the secondary data thus relates to phonetic or phonological, morphological, syntactic, or semantic properties. The encoding of these properties requires, e.g., simple attribute-value pairs (e.g. for morphological features), trees (syntax), undirected relations or pointers (co-reference).

Metadata represents another type of secondary data: information that relates to a speech or text sample as a whole and, e.g., encodes the date of recording, information about the speaker or author (sex, age, etc.). Other metadata refers to the language of the sample, in the form of typological information (e.g. ergative language), genealogical information, or areal data.

Finally, the *questionnaire* also represents a kind of annotation. The questionnaire consists of pairs of stimuli (e.g. questions, pictures, or films that are used to trigger speech) and data elicited by these stimuli. These pairs are organized in a hierarchical manner, i.e., there are more general and more specific questions, questions that presuppose other questions, etc.

2.2 Requirements

The general objective for the ANNIS effort is to provide a common database for the data collected and annotated by the individual projects. This database has to serve as long-term data storage and, at the same time, offer convenient access to the data, through search facilities and a graphical user interface for display. The research scenario characterized above places different types of demands on this database, which we briefly describe in this section.

Standard formats In order to promote convergence of the annotations performed by different projects and researchers, a common standardized annotation format is of great importance. Therefore, SFB-wide working groups are defining an *SFB Annotation Standard* with tagsets and annotation guidelines for morpho-syntax, prosody, semantics/pragmatics, and information structure.

Moreover, we are developing a common standardized representation format, the *SFB Encoding Standard*. This format represents the data and their annotations in a uniform way and allows for stating constraints on the content of the annotation. It thus facilitates the comparison of different tagsets (Which tags are used by all projects? Which tag occurs in one type of data only? Etc.). Moreover, it allows for consistency checks (only predefined tags are allowed) and completeness checks (certain annotation levels are to be annotated obligatorily).[3]

---

[3] The SFB *Annotation* Standard defines the *tag sets* that constrain the 'content' of the sec-

Further, the data of the SFB will step by step be made available to the research community. To facilitate data exchange and reuse, world-wide standard formats have to be supported: import and export format of the database must be based on XML, which allows for data exploitation and manipulation by many existing programs and tools—both general-purpose and linguistic tools, such as search tools, annotation tools, converters, and databases. Moreover, the import and export format should comply with standardized linguistic XML applications, i.e. specifications for XML-based representations of linguistic features (e.g. TEI[4], XCES[5]).

Flexibility   As mentioned above, primary as well as secondary data of the projects differ to a large extent. The database has to be sufficiently flexible to accommodate the different kinds of data. At the same time, the database should adapt to the specific needs of individual projects. For instance, sometimes intra-sentential and inter-sentential (discourse) annotation are to be combined. Hence, the database has to provide suitable visualization of both intra-sentential annotation (such as syntactic trees) and inter-sentential annotation (e.g. co-reference relations).

Querying   As studying information structure involves relating different types of information—and hence annotation—, it is important that queries to the database can easily span across different levels of annotation. Furthermore, it is important to be able to restrict the scope of queries, so that a researcher can search, for example, only the data collected by her/himself, or that assembled by a particular project, or data of a specific genre (such as spoken dialogue).

---

ondary data. The SFB *Encoding* Standard determines the format of the internal representation of primary and secondary data.

[4]   http://www.tei-c.org/P4X/
[5]   http://www.cs.vassar.edu/XCES/

Modeling of the questionnaire The database should model the structure of the questionnaire. For instance, it should allow the user to navigate from general to specific questions, to navigate from a question-answer pair in language X to the corresponding pair in language Y (whose data has been elicited on the basis of the same questionnaire).

Operability The database should be easy to operate. It should support straightforward retrieval of linguistic phenomena and an intuitive display of the primary and secondary data, so that linguists who are not experts in using databases can profit from the endeavour.

2.3   Application scenarios

The database has to be designed in such a way that it supports two rather different application scenarios. The first, henceforth called 'scenario A', is that of a centralized data repository for the SFB and beyond. Via the WWW, the data is to be made accessible to interested parties. The second, 'scenario B' is the role as research vehicle within an individual project: Data that has just been collected is annotated—maybe in a first pass rather than 'final'—and checked for consistency; first hypotheses are to be tested, which might lead to changes in the annotation; gaps in the annotation tag set might be identified. This kind of work has a clearly local, premature character and should not necessarily be executed on the 'official' central database. Instead, the system should also run on a local PC or laptop, where the projects can prepare their data until it has reached a state allowing for sharing it with others.

3   The Database

The requirements just outlined motivated the basic design decisions for the database system. In the following, we first explain its overall architecture in

somewhat more technical terms (Section 3.1).[6] Then, Section 3.2 introduces some terminology to be used in the subsequent description of the user interface (Section 3.3) and the query facility for searching data (Section 3.4).

3.1 Architecture

ANNIS is a web application that is accessed with standard web browsers. Technically, at the heart of processing are a Java servlet (which keeps all the data in memory), an open number of XML files providing the data, plus a number of DTDs, configuration files, and resources.

In addition to the requirements from the perspective of the linguistics researcher, there are a number of technical factors influencing the design. ANNIS should be:

- widely and easily accessible,

- fast with regard to display and searching,

- open with regard to integration of data from heterogeneous sources and, at the same time, supportive of our aim to create a standard format,

- open with regard to passing data on to external applications and uses,

- portable across the boundaries of operating systems, and

- configurable with regard to interface language and look and feel.

In order to comply with these goals, ANNIS was designed around the following main decisions.

---

[6] Readers who are not interested too much in technical details might want to skip this section.

Web-based Being a web application, ANNIS fulfills the criterion of universal, easy accessibility. Prerequisites on the client side are modest, as (for the most part) no special plug-ins are required. Instead, the implementation uses only HTML, CSS and JavaScript.

RAM based ANNIS is a database-backed web application. Standard usage of the term 'database' is somewhat misleading, however, since there is no genuine DBMS being used. Rather, the application reads its data from files at startup and keeps them completely in memory during a session. This was motivated by the criterion of speed; in particular, query execution profits a lot from ANNIS being memory-based. The ANNIS query language allows the construction of complex queries, employing regular expressions, grouping, disjunction, conjunction, negation, constraints on relations between nodes within trees, etc., which for an SQL processor would be expensive to analyse and execute, memory-consuming in the case of complex joins, and therefore running rather slow.

A potential reason for using a DBMS might be the ease with which data can be added, changed and deleted at runtime. However, in our application scenario A (with a centralized data repository, cf. Section 2.3), the data will be relatively stable (annotators move it from their PC to the main database only when the work is considered finished). Still, to keep track of changes, ANNIS provides an incremental update component that detects added, modified and missing files and updates the data in memory accordingly. In application scenario B (with local installation), where data change is indeed an issue, the local database can be expected to be quite small so that speed problems are very unlikely.

Dynamic importer plugin At present, data formatted according to seven different XML document type definitions (*inter alia,* stemming from the annota-

tion tools EXMARaLDA[7] and MMAX[8], and the TIGER[9] syntax annotations) can be imported into the ANNIS system. Since formats are undergoing changes and new formats are entering the scene, special care was taken to ease the process of integrating new importers. Even though we consider the development of a common XML format as an important objective for the SFB (see above), import facilities nonetheless play an important role when ANNIS is distributed to other interested parties. Therefore, ANNIS was built in such a way that adding or replacing a data importer requires no recompilation of the system as a whole. It suffices to add the new or modified Java class side by side to the other classes making up the system. It is even possible to do so in the midst of an ANNIS session: importers can be plugged in at runtime.

Export and conversion ANNIS provides several ways to export data, allowing for inspecting the data in its XML form and for externally using it in other applications. In particular, the XML data may be shown in the browser (optionally converted to an HTML representation of the data), downloaded, or sent to an email address, or deposited in a directory on the ANNIS server, optionally zip compressed.

Data can be exported in its original format, or be converted to the SFB Encoding Standard format, which we are developing as a general representation that abstracts over the peculiarities of the various annotation tool formats. At the moment, though, the SFB standard format can only be imported to ANNIS; the export module will follow.

Pure Java The use of pure Java for all server-side machinery allows ANNIS to run on all platforms providing a Java virtual machine and a Java servlet en-

---

[7] `http://www.rrz.uni-hamburg.de/exmaralda/`
[8] `http://www.eml-research.de/english/research/nlp/download/`
     `index.php`
[9] `http://www.ims.uni-stuttgart.de/projekte/TIGER/`

gine. So far, ANNIS has been installed on Windows NT, Windows XP, Mac OS X and Linux, in each case under an Apache Tomcat web server running in standalone mode.

Localization and adaptation At present, the user interface can be configured to run in English or German mode. The localization for other languages would pose no problem. Users may adapt the appearance of ANNIS in a number of ways, e.g. with regard to screen size, tool tips, and the like. Administrators can in addition control colors and other elements of style.

3.2 Concepts and notions

In the next sections, we address visualization and querying of data within AN-NIS. To ease reference to the data and concepts of ANNIS, we now introduce some notions and illustrate them by example annotations. The example text is annotated by part of speech (POS), cognitive status (COGN-ST), topichood (TOPIC), see Figure 1, and syntax, see Figure 2. The tags used there will be explained below.

Primary data Primary data is the source data, i.e. the text (or speech) that is to be annotated by linguistic data. The primary data in the example in Figure 1 is *Eier-Produzenten aus der ganzen Republik machen...* ('Egg producers from all over the republic make...').

Secondary data Secondary data consists of the linguistic data that is attached to primary data. For instance, the part-of-speech annotation in Figure 1 represents secondary data.

TOPIC	aboutness-topic						
COGN-ST	inferrable						
POS	NN	APPR	ART	ADJA	NN	VVFIN	
Text	Eier-Produzenten	aus	der	ganzen	Republik	machen	…

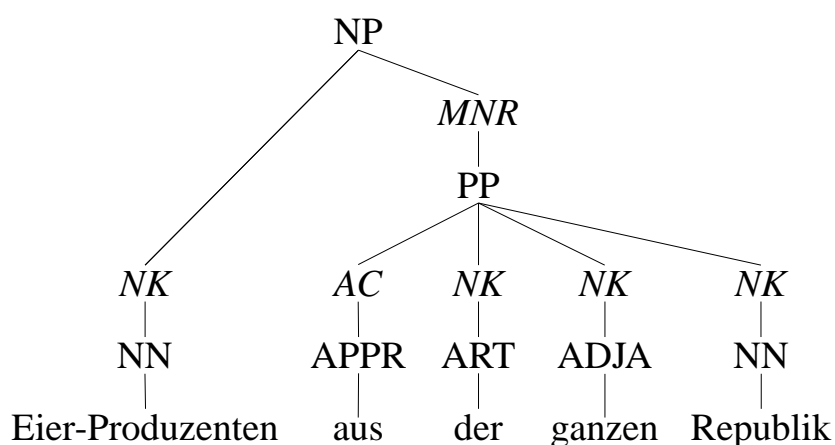Figure 1: Example annotation, encoding topichood, cognitive status, and part of speech

Figure 2: Example annotation, encoding part of speech, syntactic categories, and grammatical functions (TIGER syntax; functions are set in italics)

Annotation level Annotations are grouped according to linguistic domains, which correspond to annotation levels, e.g. part-of-speech or information-structural annotation levels.

Complex linguistic domains may be broken into smaller levels. For instance, information-structural properties can be represented by different annotation levels, such as cognitive status and topichood as in Figure 1.

Competing analyses of the same domain are considered distinct annotation levels. For instance, there can be an STTS[10] POS annotation level (i.e., the analyses comply with the STTS annotation guidelines) vs. an SFB POS

---

[10] `http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/ stts-1999.ps.gz`

annotation level (with analyses according to the SFB Annotation Standard for part-of-speech annotation).

Each annotation level is characterized by a specific tagset.

Tagset A tagset is the set of attribute-value pairs (= tags) that are admissible at a specific annotation level. For instance, the part-of-speech annotation level can specify STTS as its tagset. STTS makes use of only one attribute, "pos", with 51 different values: "NN" marks common nouns, "APPR" prepositions, etc. Accordingly, an STTS-compliant attribute-value pair is "pos=NN".

Syntactic tagsets often use two attributes, "cat", which encodes the syntactic category, and "func", encoding the grammatical function. Admissible values for the attribute "cat" might be "NP", "PP", etc., and "NK" (noun kernel), "MNR" (modifier of a noun, postnominal ('right')), "AC" (adpositional case marker) for the attribute "func", cf. Figure 2.

Tag An attribute-value pair is called 'tag', e.g. "pos=NN", "cat=NP".

(Atomic) annotation These are the elementary units of any annotation. An atomic annotation consists of a tag that is attached to a segment, i.e. to a piece of primary data (e.g. text) or secondary data (a sequence of atomic annotations).

(i) An atomic annotation can consist of an attribute-value pair that is attached to a piece of primary data. For instance, the annotation "pos=NN" in Figure 1 is attached to the token *Eier-Produzenten*, the annotation "cognitive-status=inferrable" is attached to a sequence of tokens, *Eier-Produzenten aus der ganzen Republik*.[11] Put differently, "pos=NN" is one of the atomic annotations

---

[11] Technically speaking, part-of-speech annotations are not attached directly to primary data in our implementation. We define characters as the basic units, i.e., atomic annotations of type "char" mark single characters. Next, atomic annotations of type "tok" refer to the basic "char" annotations. "pos" annotations are then attached to "tok" annotations; "pos" (and "tok") annotations are therefore atomic annotations of type (ii) rather than (i).

TOPIC	aboutness-topic						
	is-domain						
Text	Eier-Produzenten	aus	der	ganzen	Republik	machen	. . .

Figure 3: Example annotation, encoding the annotation level of topichood (TOPIC) by two attributes displayed on two tiers

of *Eier-Produzenten*, and "cognitive-status=inferrable" is one of the atomic annotations of *Eier-Produzenten aus der ganzen Republik.*

(ii) An atomic annotation can consist of an attribute-value pair that is (recursively) attached to one or more atomic annotations (this is needed for the encoding of hierarchical structures such as trees). For instance, the atomic annotation "func=NK" in Figure 2 is attached to the atomic annotation "pos=NN". The atomic annotation "cat=NP" is attached to a sequence of atomic annotations, "func=NK" and "func=MNR".

Segment   A segment defines a sequence of primary or secondary data: a piece of text (a sequence of characters or tokens), or a sequence of atomic annotations.

Instantiated annotation level   The set of all atomic annotations belonging to an annotation level is called 'instantiated annotation level'. That is, an instantiated annotation level consists of all attribute-value pairs that are actually used in the annotation—as opposed to the tagset, which defines the range of all attribute-value pairs that could be used.

Annotation layer   An annotation layer is the graphical display of an instantiated annotation level. One annotation layer consists of one or more tiers that are stacked on top of each other. For instance, the annotation level of topichood might define two attributes: "aboutness-topic" and "is-domain", which marks general information-structural domains. The segments annotated by these at-

tributes always overlap, since any topic expression must be located within an information-structural domain. Hence, the display of the atomic annotations is spread over two tiers—one displaying the attribute-value pairs of "aboutness-topic", the other displaying "is-domain"—to make the extensions of the segments transparent, cf. Figure 3.[12]

Tier   A tier is part of an annotation layer: one line, displaying atomic annotations.

Document   A document consists of primary data plus all instantiated annotation levels that refer to this data. In our examples in Figures 1 and 2 (which are based on the same text), the text and the instantiated annotation levels of part of speech, cognitive status, topichood, and syntax form a document.

Corpus   A set of documents is a corpus. Corpora can be defined according to criteria such as 'documents with the same object language', 'documents annotated with TIGER syntax', 'documents of the SFB project X', etc.

3.3   Visualization

3.3.1   Tier model

The basic metaphor of visualizing the annotated data in ANNIS is that of a *tier set*. The data window thus consists of a single line of primary text at the bottom, and a variety of annotation layers on top of it. For illustration, Figure 5 below provides a screenshot. Each annotation layer can use its own segmentation of the primary text (with the character being the minimal unit). Browsing through the text for the user means 'horizontal scrolling', for which ANNIS supplies

---

[12] Instead of distributing the information over multiple tiers, other visual means can be exploited, e.g. bubbles emerging on mouse-over; see Section 4.1.

functions to move the text (and its annotations) forward or backward, character-wise, or in jumps with adjustable lengths. This display mode largely mirrors that of tier-based annotation tools such as EXMARaLDA[13] or Praat[14], and users who are experienced with such tools should get used to ANNIS quite quickly. In addition to the annotation window, the complete source text is displayed at the top of the page, with the portion currently shown in the main annotation window being underlined, so that the current position in the complete document is always transparent.

3.3.2   The role of trees

Opting for a tier-based mode to structure and display the data entails that trees are not the primary vehicle for conveying information. Trees can of course be shown in tiers, but this is not the most natural way to present them (cf. Figure 7 and the discussion in Section 4.3.2). The decision to center the data around a tier-model rather than a tree model followed from the primary purpose of the project: Investigating information structure by seeking correlations between quite different kinds of annotations is easier when the annotation and its visualization makes as little a commitment on structure as necessary—and tiers are the most versatile scheme in this respect.

However, ANNIS offers the possibility to associate images with database entries, in which case a hyperlink is given as part of the data. Pre-stored images of tree structures can be accessed this way, for instance using SVG-files that can be exported from TIGERSearch[15] and displayed in the web browser by the Adobe SVG interpreter. In the same fashion, sound files can be added to the data.

---

[13] `http://www.rrz.uni-hamburg.de/exmaralda/`
[14] `http://www.praat.org/`
[15] `http://www.ims.uni-stuttgart.de/projekte/TIGER/TIGERSearch/`

3.3.3 Displaying multiple layers

The most interesting data for the purposes of the SFB is data annotated with different types of information. For example, the *PCC10* annotations, which will be explained in Section 4, use six annotation levels. When inspecting the data, not each layer will be of relevance for each purpose. Thus, layers can be clicked away individually so that the users can focus their attention on the type of information that is currently of interest.

When the labels of atomic annotations shown in annotation layers have to be shortened (to fit the size of the unit), the full version of the label automatically appears on mouse-over. Similarly, when viewing the tagset, extended explanations can be shown on mouse-over. These and some other features can however be configured by the user (whether they appear on mouse-over or on click, what is the window size, etc.).

3.4 Querying

Similar to visualization, querying in ANNIS is both flexible and adaptable to specific needs. It offers a rich set of search operators that can be applied to different types of data: (i) primary data (text), (ii) secondary data (annotations), and (iii) corpora (collections of annotated texts).

Text searches refer to the surface string (or the transcription of speech); for instance, one can search for specific words (e.g. *erst* 'only'). Annotations can be searched for attributes (e.g. "topic") or attribute-value pairs, including relations and pointers. Queries for corpora usually occur in combination with text or annotation queries. They allow the user to narrow down the search space by specifying an individual document or a set of documents. For instance, the query can be restricted to documents of a specific SFB project.

3.4.1   Search expressions

Wildcards, regular expressions   Basic searches relate to one word (e.g. `erst` 'only')[16] or to one atomic annotation (e.g. `cognitive-status=inferrable`). These search expressions can make use of wildcards, i.e. special characters that match any character in the string comparison. For instance, `pos=N*` matches both expressions marked as "pos=NN" (common nouns) and those marked as "pos=NE" (proper nouns). Text queries may even use regular expressions: `sag(en?|st|t)` matches surface forms like *sage* or *sagst*.

Cross-level queries   Often, queries refer to atomic annotations on different levels, e.g. in a search for an expression that is both annotated as the subject and as being inferrable. Such restrictions can be freely combined by means of the Boolean expression "&": `function=subject & cognitive-status=inferrable`. In ANNIS, these restrictions are evaluated with respect to the text that is annotated by the respective attributes. The query example is then interpreted as follows: any piece of text marked as a subject satisfies the restriction of the first conjunct, and any piece of text marked as being inferrable satisfies the second part. Combining both conditions means in ANNIS: looking for text fragments (within the text pieces) that satisfy both conjuncts simultaneously. That is, the text pieces satisfying the first and second conjunct must overlap and the overlapping part qualifies as a match.[17]

For instance, an annotation might mark an NP as the subject; suppose the NP contains an attributive adjective that is marked by contrastive focus, as illustrated in Figure 4. In this annotation, the adjective fulfills both constraints

---

[16] In this section, expressions in `typewriter` denote actual query expressions that can be typed into ANNIS. Some of the examples are slightly simplified, though.

[17] Technically speaking, the annotations that satisfy the conjuncts are part of the actual match as well.

```
              NP
            func=
           subject

     D        ADJ        N
            focus=
          contrastive
```
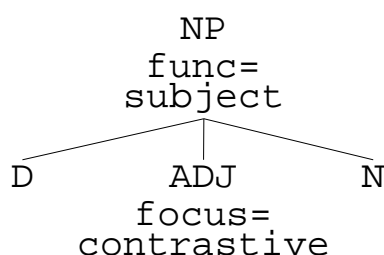
Figure 4: Focus-marked adjective within a subject NP

(being (part of) a subject and being marked by contrastive focus) and, hence, counts as a match to the query.

A stronger restriction requires that text pieces satisfying the conjuncts be identical: `function=subject _=_ focus=contrastive`. Here, the complete subject NP would have to be marked by contrastive focus.

Complex conditions The above examples illustrate the combination of restrictions by means of "&" to form complex queries. Other types of complex conditions are conditions connected by logical "|" ('or'), negated conditions, and conditions on precedence relations between expressions (e.g. an expression marked as inactive which precedes an expression marked as active: `cognitive-status=inactive .* cognitive-status=active`).

Queries for annotations in the form of a tree (e.g. syntax) can in addition refer to dominance relations, node arity (number of children), and left and right corners. For instance, `cat=NP >* cat=PP` searches for NPs that dominate PPs.

Queries across corpora As explained above, queries are evaluated by reference to the text. This means that all annotations of one text can be referenced simultaneously, even if the annotations come from different projects and are

physically part of different corpora (assuming that the same text has been anno-
tated in different corpora).

However, queries may be restricted to documents belonging to a spe-
cific corpus, by conditions on the document names: `pcc10*::cognitive-`
`status=inferrable` searches for expressions marked as inferrable in any
of the documents belonging to the corpus *PCC10* (see Section 4), i.e. *pcc10.co-*
reference, *pcc10.is.aboutness-topic*, etc.

3.4.2   Result display

Query results are delivered as a list of hits, each showing the name of the docu-
ment containing the match, the exact location of the match and the text involved
in the match. Documents on this list can be selected and are then displayed with
the matching data (text and/or annotations) highlighted. The size of the context
to be displayed along with the match can be configured by the user.

3.4.3   History, hit memory, and export

For every user, a history of the queries s/he issued is kept across ANNIS ses-
sions. In addition, users may save selected hits in their personal hit memory,
allowing search results to be revisited at a later time.

Matching documents can be exported. However, the export format of the
current version of ANNIS does not record the labels specifying those parts of
the data that actually matched the query.

4   Example: ANNIS in Action

In this section, we illustrate the operation of ANNIS with the example of the
Potsdam Commentary Corpus (Stede, 2004), a set of newspaper commentaries
that are being annotated on six different levels. In particular, we refer to *PCC10*,

a subset of ten commentaries, for which this annotation has been completed. PCC10 is annotated on the following levels:

(i) co-reference and bridging phenomena, annotated according to the guidelines proposed by Gross (2003),

(ii) information structure with aboutness topics, information focus (or 'rheme') and cognitive status (Götze, 2003),

(iii) part of speech,[18]

(iv) rhetorical relations according to RST (Mann and Thompson, 1988),

(v) connectives (Stede and Heintze, 2004), and

(vi) syntactic structure according to the TIGER treebank format (Brants et al., 2002).

4.1 Data exploration

Figure 5 shows the ANNIS user interface. The menu bar on the left is permanently visible and provides quick access to the most important functionalities of ANNIS, with a search window allowing for formulating corpus queries and navigating in the query history. The workspace on the right is the 'dynamic' part of ANNIS and is used for the various navigation and visualization tasks—for instance for the inspection of the annotation of a document in PCC10.

Our annotation view consists of three components, a *navigation bar* and a *discourse view* at the top, and a *detailed annotation view* at the center.

The detailed annotation view contains a reference line with the textual representation of the primary data at the bottom and the annotations organized

---

[18] The part-of-speech annotation has been performed by the TnT tagger using the German model, see `http://www.coli.uni-sb.de/~thorsten/tnt/`

Figure 5: ANNIS user interface

according to the annotation levels above it. In our example, annotations of the levels *pcc10.is.aboutness-topic*, *pcc10.is.cognitive-status*, *pcc10.part-of-speech* and *pcc10.rst-relations* can simultaneously be explored; other, less relevant levels (*pcc10.co-reference*, *pcc10.is.information-focus*, and *pcc10.syntax-tiger*) are clicked away by means of the triangle buttons at each annotation layer.

Annotations are best inspected by moving the mouse over the annotation at the annotation tier: this causes highlighting the primary data associated to it

at the reference line. In Figure 5, the mouse is positioned over the "aboutness-topic" in the upper center, causing *Eier-Produzenten aus der ganzen Republik* to be marked. If the mouse pauses for some time over an annotation, a bubble with more detailed information shows up, in our case displaying its tag ("topic=aboutness-topic") and the numbers representing the span of associated tokens ("42..46").

The discourse view at the top helps users to integrate the data of the detailed annotation view into the larger discourse context. The data currently focused on is underlined. By clicking on a token in the discourse view, the user can shift the annotation view so that this token appears in the center.

By means of the arrow buttons in the navigation bar, we can move back and forth in the data. We may also browse through the documents in the database by the triangular arrows (to the right of the arrow buttons).

4.2 Querying

The search window in the menu bar in Figure 5 contains a multi-level query: `cat=NP & rel_type=part-whole & topic=aboutness-topic.` This expression searches for a nominal phrase ("NP"), whose referent stands in a "part-whole" relation to a previously introduced discourse entity and constitutes an "aboutness topic". After clicking the "Go"-button, ANNIS processes the query and delivers a list of the query results. Figure 6 shows one of the results of the query. In this representation, all of the matching annotation expressions in the query are marked by underlining, i.e. "part-whole", "aboutness topic" and "NP". Again, only annotation levels specified in the query are opened up. An additional button in the navigation bar allows the user to save the result for later inspection.

Figure 6: Query result

4.3 Visualization of complex data structures

Figure 6 gives us the opportunity to consider the visualization of two further data types, pointer relations and tree structures.

4.3.1 Pointers

Immediately above the annotation tier with the underlined annotation "part-whole", a pointer annotation is shown: "-> markable: 17". This specifies a pointer relation to the annotation of a tag "markable: 17" at the very left of the *pcc10.co-reference* annotation level in Figure 6.[19] Thus, the referent of *Die Betriebe im Osten* ('The factories in the east') stands in a "part-whole" relation to the referent of the expression marked by "markable: 17": *Eier-Produzenten aus der ganzen Republik* ('Egg producers from all over the republic') in the preceding sentence.

Due to the limited size of the data segment that can be inspected in the annotation view, the current visualization is of limited use, above all for pointer relations crossing larger spans of discourse. We therefore plan to extend the functionality of the discourse view with an improved visualization of pointer relations.

4.3.2 Tree structures

In Figure 7, the tier-based representation of trees in ANNIS can be compared to conventional tree representation. The upper part reproduces a small portion of the syntactic annotation of Figure 6, and the lower part shows the corresponding tree.

---

[19] The segment that is annotated by the tag "markable: 17" only displays the number "17" in the annotation level of *pcc10.co-reference*. This segment spans the text fragment *Eier-Produzenten aus der ganzen Republik*.
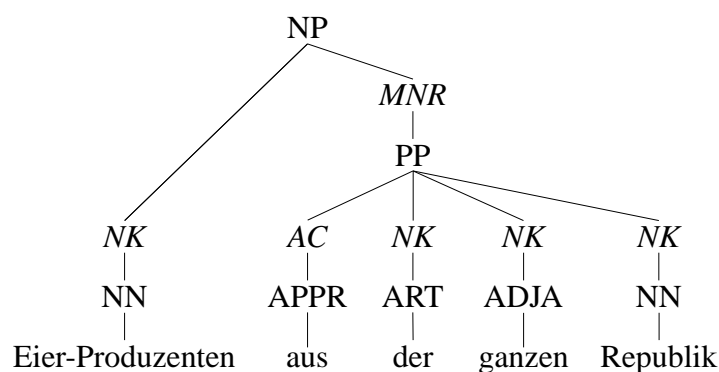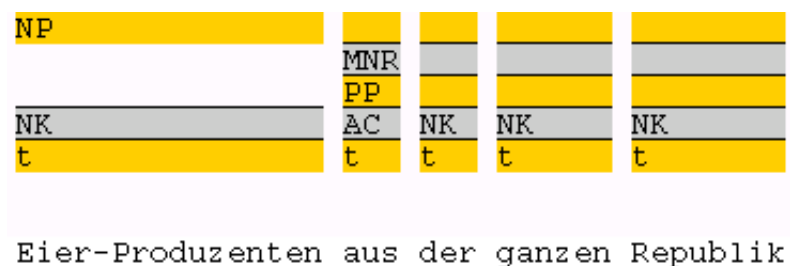
Figure 7: Syntactic annotation represented by ANNIS tiers vs. a tree

Starting from the topmost tier, tiers representing tree nodes (syntactic categories) and those representing tree edges (grammatical functions) alternate. The node "NP" directly dominates the pre-terminal of *Eier-Produzenten* via an edge "NK" (for noun kernel modifier) and the node of cat "PP" (prepositional phrase) via an edge "MNR" (for postnominal modifier). "PP" in turn directly dominates the pre-terminals of *aus der ganzen Republik* via edge labels "AC" (adpositional case marker) and "NK", respectively.

5 Evaluation

We tested the prototype in the two application scenarios presented in Section 2.3. In scenario A, ANNIS offers its services via the internet, running on a web server with a Pentium IV 2,4 GHz CPU and 3 GB memory. In scenario B, the application is run in standalone mode on a single computer, typical for

linguists wishing to work independently, for instance during field studies. We used a mobile computer with rather low hardware capabilities (Pentium III 650 MHz CPU, 256 MB RAM) for this scenario.

During the evaluation, we focused on two very general aspects: (i) the amount of data that can simultaneously be loaded into ANNIS and (ii) the querying capabilities.

5.1 Data

Due to the RAM-based approach, the amount of data that can be loaded into ANNIS depends on the memory capacities of the hosting machine. While the whole TIGER Corpus (Brants et al., 2002) with more than 40.000 syntactically annotated newspaper sentences can be loaded onto the web server (scenario A), the 1.4 GB of RAM required for this go beyond the capacities of the hardware in scenario B.

We therefore designed two data sets—L(arge) and S(mall)—for this evaluation. Both contain the Potsdam Commentary Corpus of 173 RST-annotated newspaper commentaries and the richly annotated subset PCC10. The sets differ with respect to the number of TIGER sentences they include: the former comprises the whole TIGER Corpus, the latter a subset of 1.000 sentences. Thus, data set L contains approximately 42.200 sentences, and data set S contains 3.200 sentences.[20]

With data set S, the upper limit of the amount of data fed into ANNIS is reached for the laptop. On the web server, data set L occupies 1.4 GB of RAM—even here, the limits of the hardware become relevant.

---

[20] This results in the following number of Java annotation objects in ANNIS: for L(arge): 3.369.930 objects, for S(mall): 146.505 objects.

5.2   Querying

In addition to flexible query facilities (cf. Section 3.4), ANNIS aims at providing a fast search. Beside keeping the data to be searched completely in memory, ANNIS includes running searching and result delivery in parallel threads: whenever a document is finished with, results found in it are immediately sent to the client—the user can explore the results while more results are still searched for. We therefore measured both the time until the emergence of a first result and the time needed for providing the complete list of results.

Data sets   Since data set L cannot be loaded onto the laptop, it was queried on the web server only. Data set S was tested both on the server and the laptop, enabling statements about the performance behaviour depending on the corpus size.

Example queries and evaluation method   A small set of queries of different complexity was designed: Query Q1 queries anaphoric expressions as simple attribute-value pairs; Q2 searches for expressions marked as "anaphoric" and "subject", searching across different annotation levels. Finally, Q3 exemplifies a query on hierarchical structures: it queries sentences with a subject nominal phrase that directly dominates a prepositional phrase. [21]

The queries were posed in standard web browsers and the time needed for presenting a first result and the complete list of matches was taken. Thus we did not measure the performance of the search engine alone, but the performance of ANNIS as a whole, including the construction of an HTML representation of the hit list.

---

[21] The queries have the following form:
```
Q1: rel_type=anaphoric
Q2: rel_type=anaphoric & rel=SB & #1_=_#2
Q3: cat=S & cat=NP & cat=PP & #1>SB#2 & #2>#3
```

Scenario	Query	First match (in sec.)	Completed Search (in sec.)	Hits
A (web-server)	Q1	0.5	0.5	70
	Q2	0.5	0.5	5
	Q3	2	2	130
B (mobile computer)	Q1	4	17	70
	Q2	18	18	5
	Q3	20	91	130

Figure 8: Query performance with data set S(mall)

Scenario	Query	First match (in sec.)	Completed Search (in sec.)	Hits
A (web-server)	Q1	5	8	70
	Q2	8	8	5
	Q3	2	34	4985

Figure 9: Query performance with data set L(arge)

Results and discussion   The results for querying data set S and L (given in Figures 8 and 9, respectively) show that the overall performance of the ANNIS prototype has still to be improved, particularly with respect to research scenario B, the mobile computer. Here, more complex queries such as Q3 require unacceptable processing times.

However, the strategy of an incremental presentation of query results pays off: with both data sets the first match for Q1 and Q3 is given rather quickly, even if the complete search is time-consuming.[22]

The results also illustrate the expected fact that performance of ANNIS is dependent (i) on the size of the corpus and (ii) the hardware capabilities. On the web server, queries Q1 and Q2 need considerably more processing time with the data set L than with set S. Figure 8 illustrates the difference between

---

[22] Results are currently presented document-wise. Since all hits of Q2 are part of the same document, the values 'First match' and 'Completed search' do not differ.

both research scenarios (with different hardware conditions) regarding the query performance: even the processing time for the simple query Q1 differs considerably.

Of course, these first results of a rather shallow testing cannot substitute for an in-depth study of the querying capabilities of the ANNIS search engine, which is planned to be undertaken in the near future.

6   Related Work

Current corpus exploration and query tools do not fulfill all of the needs of the SFB, as presented in Section 2. In this section, we discuss a selection of tools, concentrating on (i) web-based interfaces and (ii) query tools for complex, richly annotated data, and show how they relate to ANNIS.

6.1   Web-based interfaces

Web-based interfaces provide the quickest and easiest access to large amounts of language data and are invaluable tools for linguistic research based on corpora. Simple search facilities allow for querying the data, which usually consists of tokenized text, rarely accompanied by further levels of annotation such as part-of-speech or lemma. Search results are usually presented as plain text or as key word in context, *KWIC*. Prototypical examples are COSMAS II[23] and the online web demos of *Digitales Wörterbuch der deutschen Sprache*[24] and BNC[25].

A tool that is similar to ANNIS by providing access to very heterogeneous data and annotations is the interface of TUSNELDA ('Tübingen collection of

---

[23] IDS Mannheim, `http://www.ids-mannheim.de/cosmas2/`
[24] Berlin-Brandenburgische Akademie der Wissenschaften, `http://www.dwds.de/pages/pages_woebu/dwds_woebu_rech.htm`
[25] British National Corpus, `http://sara.natcorp.ox.ac.uk/lookup.html`

reusable, empirical, linguistic data structures')[26]. Beside searching for pure text, TUSNELDA allows the user to specify complex queries in the standardized query language XML QUERY (XQUERY)[27], which is applied to the XML-based annotations. The results of a query are shown as text (for text searches) or as XML representations (for queries on annotations). Display of the XML encoding suffices in many cases, since TUSNELDA annotations rarely cover more than one annotation level—in contrast to our research scenario.

Using XML QUERY has several advantages: Being a standardized language, it is already familiar to at least some users; the format is supported by other tools; and it is a very powerful language. Of course, using XML QUERY requires knowledge of the XML encoding of the annotation.

6.2   Query tools for complex data

In recent years, a number of tools that allow for querying and visualizing more complex annotations have been developed. These include tools for querying trees or graphs and search tools for corpora with multi-level annotation.

Trees/graphs   Examples of tree and graph query tools are VIQTORYA (Steiner and Kallmeyer, 2002), TIGERSearch[28], and Netgraph[29]. These tools enable the user to query hierarchical structures and complex relations. Moreover, they include graphical interfaces to improve operability by non-experts and casual users. These interfaces allow the user to compose a query by mouse clicks and simple menu choices. For instance, attribute-value specifications can be selected from a menu which lists all admissible attribute-value pairs. Query results are visualized as trees or graphs.

---

[26] `http://www.sfb441.uni-tuebingen.de/tusnelda-online.html`
[27] `http://www.w3.org/XML/Query`
[28] `http://www.ims.uni-stuttgart.de/projekte/TIGER/TIGERSearch/`
[29] `http://quest.ms.mff.cuni.cz/netgraph/`

However, these tools focus on sentence-based annotations of syntactic structures. That is, inter-sentential queries cannot be posed, and conflicting hierarchies (such as diverging segmentation of primary data by different annotation levels) are not accounted for.

Multi-level annotation   A tool that was developed for multi-level annotation is NXTSearch[30]. It is a highly flexible tool in that it can be applied both to time-aligned and hierarchical corpora (Heid et al., 2004). Furthermore, it allows for cross-level queries and accounts for intersecting hierarchical annotations.

NXTSearch thus offers many of the functionalities that ANNIS aims to supply. Nevertheless it does neither provide the means for visualizing and querying the annotation in a user-friendly way, nor is it accessible via the internet.

ANNIS aims at combining the advantages of the presented systems. As a web-based interface, it provides easy and quick access to linguistic data via the internet. Future development of ANNIS will profit from experiences in the user-friendly design of tools such as TIGERSearch, eventually arriving at a tool that can be easily used by non-experts. Similarly, ANNIS will build upon and continue work on multi-level and cross-level querying of tools such as NXTSearch.

7   Summary and Future Directions

We have characterized the application scenario for the ANNIS linguistic database, explained the ensuing design decisions, and described the present state of the implementation. This first version is now ready for use within the SFB and will be further developed in accordance with users' experiences. Specifically, we plan to undertake usability studies regarding both the query facilities and the visualization scheme used in the present implementation. We expect that these two topics are the central ones for further improving the system.

---

[30] `http://www.ims.uni-stuttgart.de/projekte/nite/manual/`

For querying, an option to consider is providing two different ways of accessing data: a formal query language that allows experienced users to quickly construct the expression they are interested in, and a more user-friendly one for inexperienced users, which might offer graphical options (like in TIGERSearch) and interactive help facilities. The two user groups have very different requirements, so that providing tailored access languages seems appropriate.

As for visualization, a better way of displaying trees should be integrated. Similarly, provisions have to be made to display discourse-related annotations more effectively. Co-reference information, for instance, could be shown by colouring the co-referring expressions in the discourse view (as in the MMAX annotation tool).

Within the SFB, various working groups are developing standardized tag sets and annotation guidelines (as discussed in Section 2.2). Step by step, these will be integrated into ANNIS, with the annotation guidelines made available so that users can interpret annotations that are not their own.

At least in the first round of data annotation, it might become necessary to modify the SFB questionnaire or annotation guidelines and adapt them to unforeseen data. ANNIS should thus provide a suitable way of handling data that is annotated according to different versions of the questionnaire or guidelines.

Also, some further kinds of data have to be integrated into the database:

- The questionnaire mentioned in Section 2.2 should be mapped to ANNIS so that answers can be looked for in the context of their questions; also, the hierarchical structure of the questionnaire should be preserved.

- Speech data at the moment is 'integrated' only by a hyperlink to a sound file, which might not be sufficient in the long term.

- When data in many languages is added to ANNIS, it becomes relevant to add typological information, which could then be used in the queries.

On the technical side, an important step will be adding a database to the system for application scenario A (with a centralized data repository), to ensure that ANNIS be ready to hold larger amounts of data than is possible in the present RAM-based version. Furthermore, *metadata* has to be systematically integrated into the data structures, possibly with ramifications for the query language (e.g., provide the ability to search data that originated before a specific date). Once again, existing standards such as TEI, IMDI[31], OLAC[32] will inform the design decisions.

Bibliography

Sabine Brants, Stefanie Dipper, Silvia Hansen, Wolfgang Lezius, and George Smith. The TIGER treebank. In *Proceedings of the Workshop on Treebanks and Linguistic Theories*, pages 24–41, Sozopol, 2002.

Stefanie Dipper, Michael Götze, and Manfred Stede. Simple annotation tools for complex annotation tasks: an evaluation. In *Proceedings of the LREC Workshop on XML-based Richly Annotated Corpora*, pages 54–62, Lisbon, 2004.

Michael Götze. Zur Annotation von Informationsstruktur, 2003. Diploma thesis, Universität Potsdam, Institut für Linguistik.

Juliane Gross. Algorithmen zur Behandlung von Anaphora in Zeitungskommentaren, 2003. Diploma thesis, Technische Universität Berlin, Fakultät für Elektrotechnik und Informatik.

Ulrich Heid, Holger Voormann, Jan-Torsten Milde, Ulrike Gut, Katrin Erk, and Sebastian Padó. Querying both time-aligned and hierarchical corpora with NXT Search. In *Proceedings of the Forth International Conference on Language Resources and Evaluation (LREC 2004)*, pages 1455–1458, Lisbon, 2004.

---

[31] `http://www.mpi.nl/IMDI/`
[32] `http://www.language-archives.org/`

William C. Mann and Sandra A. Thompson. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3):243–281, 1988.

Manfred Stede. The Potsdam Commentary Corpus. In *Proceedings of the ACL Workshop on Discourse Annotation*, pages 96–102, Barcelona, 2004.

Manfred Stede and Silvan Heintze. Machine-assisted rhetorical structure annotation. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING)*, pages 425–431, Geneva, 2004.

Ilona Steiner and Laura Kallmeyer. VIQTORYA—a visual query tool for syntactically annotated corpora. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002)*, pages 1704–1711, Las Palmas, 2002.

Stefanie Dipper, Michael Götze and Manfred Stede
Universität Potsdam
SFB 632, Institut für Linguistik
Postfach 601553
14415 Potsdam
Germany
{dipper,goetze,stede}@ling.uni-potsdam.de

Tillmann Wegst
Max-Braun-Str. 36
D-66123 Saarbrücken
Germany
post@tillmann-wegst.de