Out-of-focus Encoding in Gur and Kwa[*]

Ines Fiedler and Anne Schwarz

Humboldt-Universität zu Berlin

This paper investigates the structural properties of morphosyntactically marked focus constructions, focussing on the often neglected non-focal sentence part in African tone languages. Based on new empirical evidence from five Gur and Kwa languages, we claim that these focus expressions have to be analysed as biclausal constructions even though they do not represent clefts containing restrictive relative clauses. First, we relativize the partly overgeneralized assumptions about structural correspondences between the out-of-focus part and relative clauses, and second, we show that our data do in fact support the hypothesis of a clause coordinating pattern as present in clause sequences in narration. It is argued that we deal with a non-accidental, systematic feature and that grammaticalization may conceal such basic narrative structures.

Keywords: ex-situ *focus, focus marker, relative clause, conjunction, grammaticalization*

1 Introduction

This paper deals with a phenomenon concerning marked *ex-situ* focus constructions which is known from several West African languages, among them Hausa, Fulfulde and others, namely the existence of structural parallels of the out-of-focus part of these constructions with relative (and other subordinated) clauses and partly also with narrative clause types. In Hausa for example, there are two morphosyntactical codings for the perfective and the

---

imperfective aspect: a canonical paradigm A and a paradigm B which is not only found in focus constructions, but also in relative clauses (henceforth RC) and – with respect to the perfective aspect – in narrative clauses (henceforth NC). This second marked paradigm is often called the "relative" form of the respective tense/aspect and its distribution has been discussed by some authors (cf. Bearth 1993, Frajzyngier 2004).

Not so well-known up to now is the fact that similar phenomena, comprising relative and/or narrative structures in pragmatically and linguistically marked sentences, do also appear in languages of the Gur and Kwa group studied by us. In these *ex-situ* focus constructions, a focused nominal constituent takes the sentence initial position.[1] We will present their structural features in relation to the language-specific relative and narrative clause types and discuss the implications of our findings from diachronic and comparative perspectives.

Our language sample consists of five Ghanaian languages which we have been investigating in the field in 2004. Its Gur part consists of the two languages Buli and Dagbani which belong to different subgroups of the Oti-Volta branch and which are spoken in the Northern area of Ghana. The three Kwa languages considered are also of different subgroups and comprise the Inland dialect of Ewe (Gbe), the Asante dialect of Akan (Potou Tano) and the Togo mountains language Lelemi (*na*-Togo). The status of the Togo remnant or Togo mountains languages as belonging to Gur or Kwa seems however still under discussion according to Rongier 1997 (cited in Blench 2001).

---

[1] For reasons of space, we will use the term focus constituent here also in those cases in which only part of the clause constituent is focal and we will discuss in this paper only affirmative constructions. In-situ focus phenomena are not considered here at all.
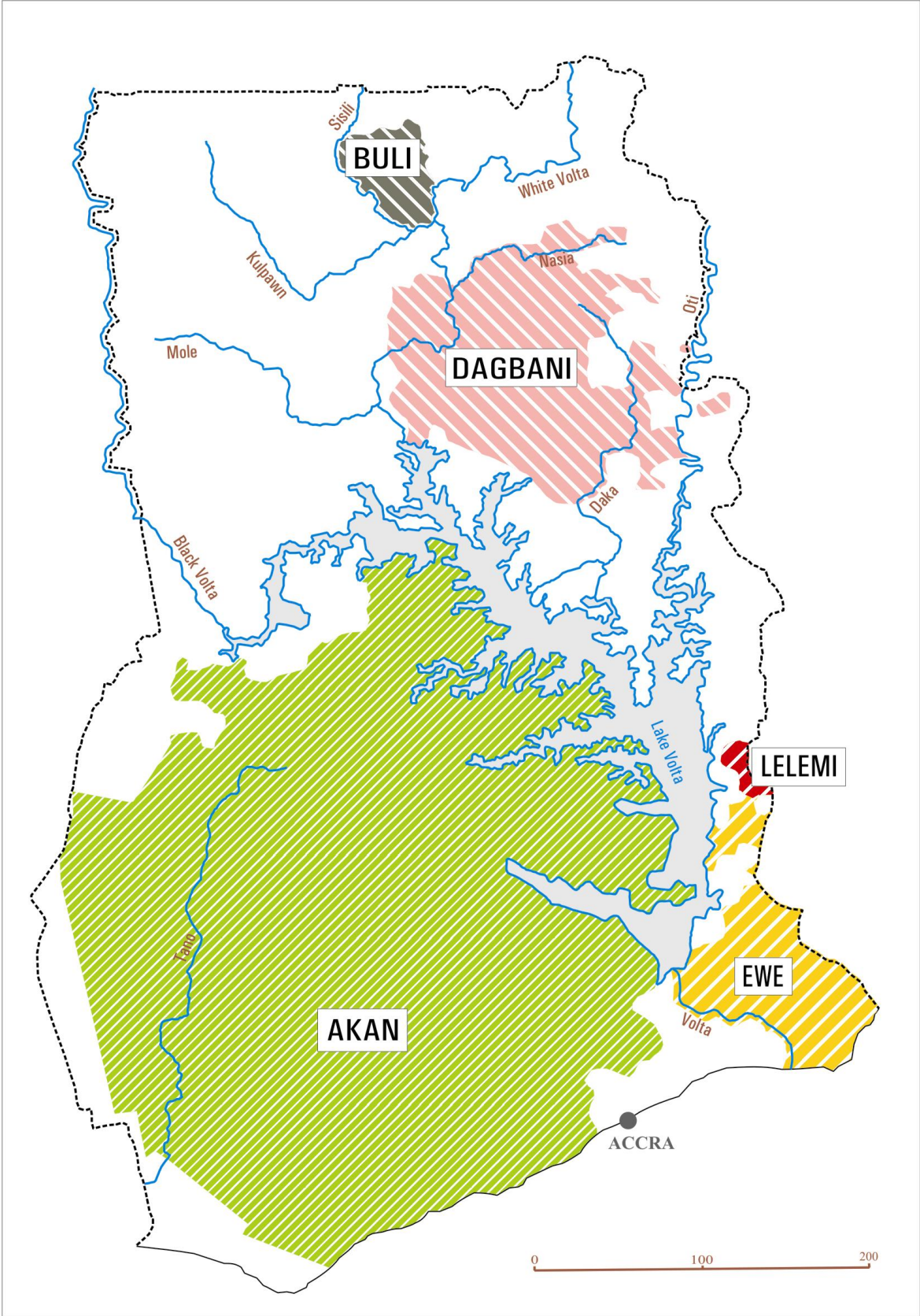
Illustration 1: Ghana Map showing our language sample

The structure of the paper is as follows: In part 2, we first present evidence for three structural characteristics that are recurring in the *ex-situ* focus constructions of almost all five languages, starting with the asymmetry between *ex-situ* subject und non-subject focus constructions (henceforth SF resp. NSF) (2.1), going on to relative-like features of these constructions (2.2) up to parallels with narrative clauses (2.3). In part 3, we first give an overview of the constructions' distribution (3.1) before we discuss the interpretation and the degree of grammaticalization of the narrative structures in each of the studied languages, and describe our findings comparatively in the conclusion (4.).

2   Structural Features of *ex-situ* Focus Constructions

In this chapter, we will demonstrate three observations concerning the structure of morphosyntactically marked focus constructions.[2]

2.1   SF vs. NSF Asymmetry

There is a constant structural asymmetry between SF and NSF constructions. This asymmetry shows up in several ways in the selected languages. Formally, we don't find the same degree of asymmetry in all the languages considered here. We will start with cases that are less obvious at first sight, and go on further to languages showing a full range of this asymmetry.

2.1.1 Ewe

The characteristics of *ex-situ* focus constructions in Ewe are as follows: First, the focused element can be marked by a FM (*y)é*, which is obligatory for subject focus and optional for non-subject focus. The exact constraints for the optionality are not yet clear. Second, there are special subject pronouns for 2nd

---

[2]   Please note that the data are transcribed with surface tones and that versalia in the English translation indicate the respective focal part of the utterance. In examples providing restrictive RCs, the relativized head and the RC are underlined.

and 3rd singular person which are used only in NSF constructions, while in SF
no pronominals are preposed to the verb.

SF

(1)   ńtsú-á-é   tsɔ́-è.
        man-DEF-YE   take-o:3sg
        'The MAN took it.' (not the woman)

NSF

(2)   èdzì(-é)   wò-   ɖù.
        top(-YE)   3sg.DEP   eat
        'He was on TOP.' (i.e., He WON the game.)

2.1.2 Akan

In Akan, the construction for both SF and NSF is characterized by two features.
It obligatorily makes use of the FM *nà* and displays a so-called "link tone"
(Bearth 2002; cf. Schachter 1973 as well) at the verb in the out-of-focus part.

In SF, an expletive subject pronoun for 3rd person referents (*ὲ-*) is
characteristically used, although it might be replaced by the normal pronominal
form.[3] In NSF on the other hand, there is no general need for an object pronoun
that is coreferent with the constituent in focus, but rather the selection of the
object pronoun underlies semantic criteria. Animates require a pronoun (cf. 4a),
especially when human, inanimates do not allow it. In the absence of a
pronominal object, the "past" transitive verb in sentence-final position carries
suffix -*yε* (cf. 4b).

---

[3]   Bearth et al. (2002) describe the change of the subject pronouns as restricted to human
      referents. This distribution is not supported by our data.

SF

(3)   ɛ̀-yɛ̀   àbrɛ̀wá   nó   nà ɛ̀-díî   àdùá nò.
         3sg-COP old.woman DEF NA 3sg-eat beans DEF
         'It is the OLD WOMAN who ate the beans.'

         cf. the canonical sentence

         àbrɛ̀wá   nó   dìî   àdùá   nò.
         old.woman DEF   eat   beans DEF
         'The old woman ate the BEANS.' ~ '... woman ATE THE BEANS.'

NSF

(4) a.   nè   krámán   nà pàpá nò   súà   nò.
         3sg dog   NA man   DEF carry O:3sg
         'The man carried his DOG.'

     b.   nè   bágè   nà   ɔ̀-súà-yɛ́.
         3sg bag   NA 3sg-carry-YE
         'He carried his BAG.'

2.1.3 Lelemi

In Lelemi, the difference between subject and non-subject focus constructions lies above all in the verbal morphology. Lelemi has two sets of TMA-markers: one used in simple tenses, the other in so-called "relative" tenses (Allan 1973). Not every simple tense has a counterpart in the relative tenses.

The "relative tenses" ("relative past, relative present, relative future and relative present for verbs of state") show up in SF constructions. Unlike the simple tenses, they don't have noun class concord for the subject, and they display only one invariant form. The simple tenses, on the other hand, are used in NSF. They are formed by subject noun class concord markers, often assimilated with the following segmentally and/or tonally coded TMA morphemes, and the verb (with grammatical tone for each tense).

The morphological coding device for NSF consists of morpheme *nà* postponed to the focused constituent (cf. the homophone FM in Akan). Some of the informants treated it as obligatory, others claimed that the syntactic marking, i.e. the sentence-initial position, makes already clear that this constituent is in focus.

SF   cf. verb form in
                                                                  simple tense:

(5) a.

    ɔ̀nàabì ùmwì pɛ́ mɔ̀-dì   kùtú.   ɔ̄ɔ̄-
dì
boy   one   only REL.PRS(dyn)-eat   orange
'Only ONE boy is eating an orange.'

    b. ɔ̀nàabì n̄-fī   ùlòkúbì.   ū-fī
boy   REL.PRS(stat)-carry   girl
'A BOY was carrying a girl.'

    c. lɔ́ɔ́lì ínyɔ́   ná-sà.
                                          lɛ́-sà
lorry two   REL.PAST-meet
'TWO LORRIES COLLIDED.'[4]

NSF

(6) àkábí áwɔ̄dí (nà) ùlōkū ɔ́mɔ̀ ɔ́-dì.
beans raw   (NA) woman DEF   3sg.PAST-eat
'The woman ate RAW BEANS.'

---

[4] In Lelemi, SF and sentence focus are coded in the same way – a feature which is in fact characteristic for Gur languages.

2.1.4 Buli

Buli has a preposed affirmative FM *ká* (negative suppletive *dāā*) in SF as well as in NSF. The affirmative morpheme seems to be optional in both constructions. In SF, the focus constituent is always followed by the conjunction *lē*, while in NSF we typically find the conjunction *tè*.[5] In NSF, the occurrence of *lē* is less common, but not totally excluded.

Verb tone deviates from that in simple sentences in both focus constructions, although not in the same way. In fact, Buli has three paradigms distinguished by grammatical tone on dynamic perfective verbs in clause-final position: a canonical paradigm A in simple sentences, a paradigm B after conjunction *lē* (SF) and a paradigm C that shows up in clauses with the conjunction *tè* (NSF).[6] In both constructions, pronouns which are coreferent with the focused constituent do not occur in the out-of-focus part.

SF

(7)   (ká)   wá   lē   chèŋ.   Paradigm B, not: *tè*
       (FM) 3sg:DISJ CNJ   go
       'HE went.' ~ 'It is HE who went.' (not you)

---

[5] Both conjunctions are sometimes provided with an initial vowel (*àlē, àtè*). This vowel occurs with other clause-inital conjunctions as well as with clause-initial serialized verbs and is always correlated with a prosodic break before the clause.

[6] Paradigm B is characterized by an "instabil rising tone" (Schwarz 2004: 38) and paradigm C by an invariable low tone. Both paradigms are constituted by the absence of subject congruent grammatical tone operating in simple clauses (Paradigm A). The neutralization in paradigm C versus A shows up only with discourse participants (1st and 2nd person), while it is not evident with 3rd persons, as in ex. 7-8.

cf.   the canonical sentence

wà   chèŋ-kà   sándēm.   Paradigm A
3sg go-FM[7]   Sandema
'He went to SANDEMA..' ~ 'He WENT TO SANDEMA.'

NSF

(8) a.   (ká) sándēm   tè   wà   chèŋ.   Paradigm C, rare:
          lē
          (FM) Sandema CNJ 3sg go
          'It is SANDEMA where he went.' (not Navrongo)

     b.   john tè   mí   fɔb.
          John CNJ 1sg:DISJ slap
          'I slapped JOHN.'

2.1.5 Dagbani

Dagbani provides a strong structural asymmetry between SF and NSF. The canonical SF construction[8] contains a postponed syllabic nasal called "emphatic" by Olawsky (1999). In NSF on the other hand, the so-called FM *kà* (Olawsky 1999: 63) has to be put at the beginning of the out-of-focus part.

Similar to Buli, both focus construction types are distinguished from each other by grammatical verb tone and deviate from the verb paradigm in simple sentences. A pronominal form for a focused subject is not allowed and in general, neither focused non-subjectival verb arguments nor other focused constituents are pronominally represented in the out-of-focus part.

---

[7] Please note that in the canonical indicative sentence in which either the complement or the whole VP might be focal, FM *ká* is enclitizised to the verb (the morpheme's surface tone may change). In *ex-situ* focus constructions any postverbal position of this morpheme is completely excluded.

[8] There is another SF construction formed with post-subjectival *léé* which is however almost restricted to questions and not considered here.

SF

(9)   páɣá   má!á   m̀   bɔ́l-ɔ̀.   Paradigm B
      woman DEF   N   call-O:3sg
      'The WOMAN called him.'

cf. the canonical sentence

      ɔ̀   bɔ̀l-lá   george.   Paradigm A
      3sg call-FM[9] George
      'She called GEORGE.' ~ 'She CALLED GEORGE.'

NSF

(10) george   kó-!ó   bɔ̀lì.   Paradigm C
      George   KA-3sg call
      'She called GEORGE.'

2.2 Relative Structures

Concerning the often stated "relative" forms in *ex-situ* focus constructions, we found that relative structures, i.e. a head and a (restricting) relative clause, are not present in all of our sample languages. And if they exist in *ex-situ* focalization, they are not necessarily identical with the language specific prototypical RCs with restrictive reading, as demonstrated in the following.

2.2.1 Ewe

Ewe disposes of a general construction to express restrictive relative clauses. The overall features of this construction are: (i) the relative clause is introduced by a demonstrativum *sì* (standard-Ewe) or *yíkè* 'this' (dialectal variant for Inland Ewe) which takes over the function of a relative pronoun and (ii) it ends

---

[9]   In case of complement or VP focus, suffixed FM *-lá* occurs in Dagbani indicative sentences. Like Buli FM *ká* in postverbal position, this suffix is excluded in *ex-situ* focus constructions.

generally with the determiner *lá*. If the relative clause follows the main clause, the determiner is sometimes omitted.

There is a difference among relative constructions depending on the syntactical function of the antecedent: if it is the subject of the relative clause, no pronominal form is used, if it is a non-subject then we find the already mentioned special pronouns which only show up in 2nd and 3rd person singular.

Subject RC

(11) éyé tsītsītɔ   sì kɔ́ èlã̀ lā vê   lá, trɔ́ dzó ...
      CNJ elder.one   DEM take fish DET come.to DET, change return
      'And the elder one, who brought the fish, returned back ….'

Non-Subject RC

(12) a. nú   sì wò-   gblɔ̀ ná mì lá   mì wɔ̀ è.
         thing DEM 3sg.DEP say for 2pl DET 2pl do O:3sg
         'What he tells you, do it!' (Duthie 1996: 45)[10]

     b. ńtsùvǐ ɖé tɔ́   ɖé
         boy   IND stand reach
         fí   yíkè wó-   nɔ̀ bɔl ɸōḿ   lè   ɸé gôl mè.
         place DEM 3pl   HAB ball beat.PROG be.LOC place goal in
         'A boy is standing in a goal.' (lit.: 'A boy is standing at the place where they normally shoot the ball into.')

There are only minor similarities between focus and relative constructions. These concern the use of the pronominal forms: as in SF, in Subject RC no pronoun is used, whereas in NSF as well as in Non-Subject RC the dependent subject pronouns are found. The two features of RC mentioned above are not

---

[10] The interlinearization was done by the authors.

present in focus constructions, and, on the other hand, morpheme (*y*)*é* of focus constructions is absent in RCs.

2.2.2 Akan

Akan relative constructions are characterized by the use of (i) the so-called "relative pronoun" *â* and (ii) a clause-final determiner *nó*. Furthermore, the verb in the relative clause changes its tone pattern in adopting a H tone (cf. Schachter 1973, the so-called "link tone" in Bearth 2002).

Subject RC

(13) <u>àbrántìè n(o)-áà ɔ-bɔ́ɔ̀ wó nó</u>, yɛ̀ m-àdámfùɔ̀.
       boy   DEF-REL 3sg-hit O:2sg DET   COP 1sg-friend
       'The boy who hit you is my friend.'

Non-Subject RC

(14) <u>àbrántìè n(o)-áà wò-bɔ́ɔ̀ nò nò</u>, yɛ̀ m-àdámfùɔ̀.
       boy   DEF-REL 2sg-hit O:3sg DET   COP 1sg-friend
       'The boy whom you hit is my friend.'

Hence, in relativization two additional features show up compared to focus constructions. On the other hand, post-focal morpheme *nà* is missing in RC. What both constructions share in comparison to simple sentences is only the "link tone".

2.2.3 Lelemi

Lelemi displays a subject- and non-subject asymmetry in relative constructions that resembles the dichotomy in its *ex-situ* focus constructions (cf. 2.1.3) and that is based on the syntactic function of the relativized element within the RC. Common component in both types of relative constructions is the determination of the relativized element by an identifier pronominal form that consists of a noun class concord for the preceding relativized noun and morpheme -*ní*. If the subject is relativized, a "relative" TMA form, i.e. the TMA form without subject

prefix at the verb is used (cf. the simple verb form in the initial clause of 15 with the relativized below). If a non-subject constituent is relativized, the simple TMA verb form, i.e. including subject prefix, occurs.

Subject RC

(15)   èbùò   ū-ɲè.
        animal 3sg.PRS-stand
        'There is an animal there,

        ńzù   èbùò   ɔ́-ní   n̄-ɲè   vɔ̄   ɔ́-dì   ɔ̀náànjùé.
        but   animal 3sg-NI REL.PRS(stat)-stand there 3sg-COP cow
        but the kind of animal that is over there is a cow.'

Non-Subject RC

(16)   ɔ̀klàmá ɔ́ŋvɔ̀ ɔ̀ɔ̄-dù   trouzis ɔ́-ní   ɔ̄nāābì ɔ́ŋvɔ̀ ɔ́-chà
        dog   DEM 3sg.CONT-bite trousers 3sg-NI boy   DEM 3sg.PAST-wear
        'The dog bites the trousers which the boy wears.'

Apart from the use of the "relative" versus the simple TMA-forms, *ex-situ* focus constructions and RCs with restrictive meaning are not the same. It is first of all the "identifier pronoun" (noun class pronoun *–ni*) which distinguishes the restrictive RC from focalization. A further element that is typical for NSF contrary to relativization is morpheme *nà* at the beginning of the out-of-focus part.

2.2.4 Buli

Buli disposes of two structural types of RC[11] which share the following features: (i) the relativized head is provided by an indefinite noun class pronoun[12], like

---

[11] Cf. Hiraiwa 2003 for a detailed description of RCs in Buli.
[12] The indefinite pronouns refer to specific indefinite entities and are therefore translated as 'certain, some' in other contexts.

wāā(i), which can either represent the head on its own or forms a nominal compound with an initial nominal constituent, and (ii) determiner *lá* is commonly added to the end of the relative clause. This determiner is sometimes omitted in sentence-final RCs. The head internal RC type shares features with SF constructions, since it consists of preverbal conjunction *lē*. It is always used when the relativized element has subject function in the relative clause and it can be found with relativized verb objects, too[13]. The other RC is of the head-external type and not compatible with a subject as head. Structurally, this RC resembles NSF, since it contains conjunction *tè* (ex. 18b). The grammatical tone of perfective verbs deviates in both RC types from paradigm A with tonal subject agreement.[14]

Subject RC

(17)   <u>núrú-wāā   lē chèŋ lā</u> ká mí   dɔā.
        person-IND:CL CNJ go   DET COP 1sg:DISJ friend
        'The person who has left is my friend.'

Non-Subject RC

(18) a. <u>fì lē yàlì nùrú-wāā   lá</u> ká mí   dɔā.
         2sg CNJ marry person-IND:CL DET COP 1sg:DISJ friend
         'The person you married is my friend.'

     b. <u>núrú-wāā   tè fì yàlì lā</u> ká mí   dɔā.
         person-IND:CL CNJ 2sg marry DET COP 1sg:DISJ friend
         'The person you married is my friend.'

---

[13] The head-internal RC therefore has either a postnominal (relativized subject, cf. ex. 17) or a circumnominal (relativized object, cf. ex. 18a) structure.

[14] In RC with clause-final determiner *lá* there is however no "instabil rising tone", since this paradigm B pattern is restricted to perfective verbs in clause-final position (cf. clause-final *chèŋ̄* in ex. 7 versus ex. 17).

Despite the similarities in morphology (conjunctions *lē* and *tè*) and tonal neutralization (no paradigm A for dynamic verbs in perfective), both relative clause types with restrictive reading contain components which distinguish them from the *ex-situ* focus constructions: the indefinite class pronoun and the RC-final determiner.

2.2.5 Dagbani

Similar to Buli, Dagbani has two RC types at its disposal which share the following features: (i) the head is represented by an indefinite noun class pronoun (like *so* in examples (19) and (20))[15] that forms a compound with the nominal root or is used alone, and (ii) determiner *máá* (sometimes *lá*) is added to the end of the RC.

One of the RC types is restricted to cases in which the antecedent has subject function within the RC (ex. 19). Apart from the two features mentioned above, it is formed with the help of a disjunctive pronoun in the subject slot which follows the head and fulfills here the function of a relative pronoun.[16] The other RC type occurs only with non-subjects as relativized heads and makes use of post-subjectival particle *n(i)* marking also some other subordinated clauses. In this head-internal RC type the head is either retained in its postverbal slot (cf. the circumnominal ex. 20a) or it is moved to the initial position of the relative clause (cf. the postnominal ex.20b).

Subject RC

(19) <u>dò-só</u>   ŋún   !cháŋ má!á ɲέ-!lá   ń   zò.
      man-IND:CL 3sg.DISJ go   DET   COP-FM 1sg friend
      'The man who has left is my friend.'

---

[15] The indefinite pronoun consists of a stem *s[front vowel]* (the vowel occurs only in case of CV suffixes) which is provided with a class suffix. It corresponds functionally to the indefinite pronoun in Buli (cf. footnote 12), although the latter lacks the *SV* stem.

[16] According to Wilson (1963: 139), the indefinite pronoun is optional in subjectival relative clauses.

Non-Subject RC

(20) a.   <u>á-n(í)   !ŋmé dò-só   má!á</u> ɲɛ́-!lá   ń   zò.
           2sg-CNJ hit   man-IND:CL DET   COP-FM 1sg friend
           'The man whom you have hit is my friend.'

     b.   <u>dò-só   á-n(í)   !ŋmé má!á</u> ɲɛ́-!lá   ń   zò.
           man-IND:CL 2sg-CNJ hit   DET   COP-FM 1sg friend
           'The man whom you have hit is my friend.'

The comparison with the focus constructions shows no direct correspondence, but it is possible that the subordinating particle *n(i)* and the "emphatic" syllabic nasal *N* that obligatorily occurs in SF are related to each other.

2.3 Narrative Structures

Our third observation concerns the fact that there is a constant formal parallelism between *ex-situ* non-subject-focus constructions (NSF) and narrative clauses (NC), and in some of our sample languages the narrative structure is extending to SF, too. We use the term NC for clauses that encode the succession of events in realis mood and that serve to continuitively develop the main story line. Labov regards this function as follows: "Each clause … describes an event that is understood to shift reference time, i.e. it follows the event immediately preceding it, and precedes the event immediately following it." (1972, cited in Schiffrin 1994: 284).

The formal parallels show up in several ways in the selected languages and are realized by the FM, TMA morphology including tone, and special pronominal forms.

2.3.1 Ewe

Ewe shows certain structural similarities between NSF and NC, although there is no total correspondence. These are best demonstrated by the use of special subject pronouns[17] in both constructions, as can be seen in ex. (21) and (22).

Additionally, there is a similarity of the FM with a clause coordinating conjunction *éyē* 'and (then)' which is found in narrative contexts.

NSF

(21)   èdzì(-é)   wò-   ɖù.   (= ex. 2)
        top(-YE)   3sg.DEP   eat
        'He was on TOP.' (i.e., He WON the game.)

NC

(22)   And the third one .. found the way to the market …

        yá   wò-   vá   kpɔ̄   tòmátós.
        CNJ[18] 3sg.DEP come   see   tomatoes
        '… and he found tomatoes.'

2.3.2 Akan

Akan has a clausal sequential conjunction *nà* with the meaning 'and (then)' (Bearth 2002) which is identical with the FM. The verbal morphology including the "link tone" also seems to be the same in both clauses, although this is still a matter of further research.[19]

---

[17] This has already been noted by Duthie (1996: 53) and Ameka (2004: 17).

[18] The form *yá* is one dialectal variant of the conjunction *éyē* in Inland Ewe.

[19] Bearth (2002) postulates the existence of a "link tone" on the verb as well as the existence of the so-called "dependent" morpheme *ye* only in *ex-situ* focus constructions, while our own data exhibit them in other contexts, too, including sequential events with clause-initial conjunction *nà*.

NSF

(23)   àdùá   nà   ɔ̀-díì-yɛ́.
        beans NA   3sg-eat-YE
        'He ate BEANS.'

NC

(24)   máàmé nòáà àdùá,   nà   n-àdámfùò díì-yɛ́.
        Maame cook beans   CNJ   3sg-friend   eat-YE
        'Maame cooked beans and her friend ate them.'

2.3.3 Lelemi

In Lelemi, NSF and NC clauses show identical features: In both, the simple tenses are used. Furthermore, the FM is homophone with the narrative conjunction 'and' which coordinates two clauses and we suppose that it is the same morpheme. It is segmentally identical with the "relative past" tense morphem, too (cf. ex. 5c).

NSF

(25)   àkábí áwɔ̄dí (nà) ùlōkū ɔ́mɔ̀ ɔ́-dì.   (= ex. 6)
        beans raw   (NA) woman DEF   3sg.PAST-eat
        'The woman ate RAW BEANS.'

NC

(26)   'The youngest child went …'

        nà   ú-tī   ùlū   gɛ̀ɛ̄ɛ̀   ɔ́mɔ̀.
        CNJ   3sg.PAST-take road   right   DEF
        'and he took the right road.'

2.3.4 Buli

Buli, too, displays a striking parallel between NSF and NC. First, the clause initial element *tè* of the out-of-focus part of NSF and the clausal conjunction *tè*

'and' in NC[20] are completely identical. Second, the identity pattern stretches onto the verb: it bears the same grammatical tone (paradigm C) after *tè* in both functions and it differs thus from (i) the canonical paradigm A, and (ii) the paradigm B that is found after marker *lē* (i.e. mainly with SF, cf. 2.1.4.). This is true only for the group of dynamic verbs in perfective aspect as shown in the following examples. Dynamic verbs in the imperfective as well as stative verbs do not participate in paradigm C, but return to paradigm A where they are most often – although not throughout – marked as irrealis, i.e. the non-perfective verb forms tend to occur in subjunctive in focus constructions.

NSF

(27) (ká) sándēm   tè   wà   chèŋ.   (= ex. 8a)
      (FM) Sandema   CNJ   3sg   go
      'It is SANDEMA where he went.' (not Navrongo)

NC

(28) and his mother was happy with him

      tè   bà   dìg   jèntàŋā ...
      CNJ   3pl   cook soup.DEF
      'and they cooked the soup ...'

2.3.5 Dagbani

Finally, Dagbani also affirms the parallel pattern between NSF and NC convincingly. Morpheme *kà* which is following the focused constituent in NSF constructions has a clausal conjunction counterpart *kà* 'and' in narrative contexts. Furthermore, the grammatical tone of dynamic verbs in such clauses differs in the same way from the subject congruent verb tone in simple clauses,

---

[20] At the beginning of subjunctive clauses (irrealis), the conjunction is also consecutively used.

irrespective whether we deal with a real narrative context or with a focus construction.

The coding of the second clause in ex. (30b) demonstrates Olawskys (1999: 44) observation that, if the subject of the clause introduced by *kà* is coreferent with the subject of the preceding clause, it has to be elided.

NSF

(29) yílə́ má!á nì kó-!ó dì.
      house DEF in KA-3sg eat
      'In the house she ate.'

NC

(30) a. and the mother sent the youngest child

      kà bíí má!á chàŋ ...
      CNJ child DEF go
      'and the child went ...'

   b. páɣá má!á dáá-!lá peter ká !ŋmé-ò.
      woman DEF push-FM Peter CNJ hit-O:3sg
      'The woman pushed and hit Peter.'

                        not: ... ka *o ŋme-o.

2.4 Diversity and Distribution of Forms

Summarizing our observations concerning the structural features of affirmative *ex-situ* focus constructions, relative clauses and narrative clauses, we have to state that the *ex-situ* focus constructions minus the focus constituent itself (F) resemble relative clauses only to a certain extent while the structural features of narrative clauses are matched much closer. Table 1 gives an overview on the differences and parallels as they have emerged in our investigation:

Table 1

	Kwa Languages			Gur Languages	
	Ewe	**Akan**	**Lelemi**	**Buli**	**Dagbani**
SF	F *(y)é*	*(ὲyὲ)* F *nà* expl. *ὲ* link tone	F rel. tense	*(ká)* F *lē* tone B	F *N* tone B
RC₁	*sì~(yí)kὲ* DET *lá*	*â* link tone DET *nó*	CL-*ni* rel. tense	IND:CL *lē* DET *lá*	IND:CL *n(i)* DET *la~máá*
RC₂	*sì~(yí)kὲ* DEP pron. DET *lá*	--	CL-*ni* simple tense	IND:CL *tὲ* DET *lá*	IND:CL DISJ pron. DET *la~máá*
NSF	F *(y)é* DEP pron.	*(ὲyὲ)* F *nà* link tone	F *nà* simple tense	*(ká)* F *tὲ* tone C	F *kà* tone C
NC	... *éyē* DEP pron.	... *nà* link tone	... *nà* simple tense	... *tὲ* tone C	... *kà* tone C

With respect to RCs, we face considerable heterogeneity among our sample since the two Gur languages provide especially strongly divergent RC types. This is due to the additional head-internal relative clause type in these two languages which is represented by RC₁ in table 1 and which seems structurally related with the SF constructions in both languages (cf. preverbal morpheme and grammatical tone paradigm). The head-external RC type among these languages is represented in the table by RC₂, but only in Buli it also displays evident formal parallels with a focus construction, namely with NSF. Considering the relation between RCs and the non-focal part of *ex-situ* focus constructions in the Kwa languages, we again note only partial correspondences: in Ewe, dependent pronouns occur in RCs when a non-subject is antecedent as well as in NSF; in Akan, tonal changes ("link tone") pertain in RC as well as in SF and NSF; and in Lelemi, the selection of the tense form in RC and focus constructions is due to the syntactic function of the preceding relativized respectively focused element as subject or non-subject. Absent in focus constructions are however those morphological means which are characteristic for almost all restrictive RCs

throughout our language sample, i.e. the respective (relative / identifying / indefinite) pronominal forms accompanying the head and, with the exception of Lelemi, the RC-final determiners.

By integrating NCs into our considerations, it becomes evident that part of the so-called "relative" features in focus constructions are not just exclusively "relative", if at all, but should rather be analysed as structural reflections between coordinated "narrative" clauses. Contrary to the complex picture with regard to the RC pattern in focus constructions, all five Ghanaian Gur and Kwa languages considered here display in fact a very close correspondence between (N)SF and NC. With the exception of Ewe and some need for verification in Akan, we can even postulate a complete structural identy for both. We therefore conclude that the parallelism between the out-of-focus part of morpho-syntactically marked (N)SF and narrative clauses (NC) is no coincidence, but is due to a systematic "narrative" basis of the respective focus constructions.

3   Narrative Hypothesis

From the structural distribution above it is evident that the parallelism between (N)SF and NC is a systematic pattern. We propose that in fact a narrative clause constitutes the non-focal part of such *ex-situ* focus constructions and that its grammaticalization can conceal their biclausal structure. We therefore don't follow the movement hypothesis as for example suggested by Aboh (2004) and Green (1997) for Gbe and Hausa respectively. Like in the cleft hypothesis assumed for various languages (Givón 1990/2001, Schachter 1973, Heine/Reh 1984, Lambrecht 2001), our proposal considers the constructions as biclausal and adds a diachronic perspective to their synchronic analysis. Unlike in the prototypical cleft hypothesis however, we here assume a NC rather than a RC as source for the potential or the already undergone grammaticalization processes. In the following, we will argue for the validity of the narrative hypothesis for each language separately.

3.1 Ewe

Ewe focus constructions can be regarded as derived from two coordinated clauses, where the second clause is provided by an original narrative clause while the first clause is commonly represented by an NP alone. A copula form is only needed in the latter, if the focused constituent is negated, a fact which holds for all the languages considered here.

The synchronic FM *yé* shows structural similarities to the conjunction *éyē* (which is underlying /*éyé*/ [21]) 'and, and then', although it is not identical. The meaning of the conjunction's prefix *é-* hasn't been explained convincingly up to now. We assume that the FM has developed out of the conjunction. Following this grammaticalization path, one has to claim a divergent development of the synchronic conjunction and FM. The latter is in normal speech usually eroded to vowel -*é* and suffixed to the preceding NP, that is, it has become part of the initial clause.

Synchronically, Ewe displays a homophone morpheme *yé* occurring in nominal predication like 'It is a pen.' – *pên yé*. Here, its function is comparable to a copula verb.

A further structural feature supporting the narrative hypothesis in Ewe is the use of the special subject pronouns in NSF and NC. Westermann (1930: 61) mentions that they are used "in the continuation of a sentence, or closely to connect one sentence with a preceding one." It is only in subject focus constructions, where they are not required and even ungrammatical. If we assume a narrative construction underlying both, SF and NSF, then we have to state that in SF the structure of the original source is extensively eroded due to a phenomenon we will call here the "double-subject" constraint, i.e. the focused constituent cannot be followed by a coreferential pronominal subject in SF.

---

[21] Cf. Clements (1977: 172) for the tone rules changing the two phonological high tones of the conjunction.

3.2 Akan

The first hint for the validity of the narrative hypothesis in Akan is the identity of FM and narrative conjunction. Our informants treated *nà* in FC still as conjunction so that, if there is a certain degree of grammaticalization at all, as suggested by its description as FM by some authors, this could be only by a functional split in the very inceptive stage.

The first clause of the biclausal focus construction is often only represented by an NP. Alternatively, the initial clause starts with *èyè*, i.e. an expletive pronoun plus copula verb 'to be' (cf. ex. 3).

Different from Ewe, in Akan, the biclausal status of the subject focus construction is still well maintained, since the out-of-focus clause obligatorily requires a subject pronoun, which might be an expletive one (cf. 2.1.2.). The "double subject" constraint is thus not operating in this language.

3.3 Lelemi

In Lelemi, the narrative clause as part of an *ex-situ* focus construction is evident on first sight only for NSF. As shown above, the non-focal clause of NSF and the narrative clause are formally totally identical, i.e. any probably assumed grammaticalization of the narrative clause is restricted to the functional level and has no effects on the structural level. Accordingly, morpheme *nà* is in both functions considered as conjunction by us.

In SF, the conjunction is missing and "relative" tense forms are used instead of the simple tense forms. Nevertheless, we can assume such grammaticalization source in one of the relative tenses/aspects. The TMA morpheme for the "relative past" tense is high toned *ná*. We analyse it as a conglomeration of the conjunction *nà* (with inherent Low tone) plus a High tone which is born by the subject pronoun in simple past. The slot for the pronoun is not filled due to the "double subject" constraint. The high tone it bears in simple past is however retained with the former conjunction.

3.4 Buli

In Buli, the narrative hypothesis is valid for the prototypical NSF construction which is formed with conjunction *tè* and tone paradigm C. Since these features are shared by sequential clauses in narrative contexts, too, a narrative clause can be regarded as representing the non-focal part in NSF. The lack of tone paradigm C with stative verbs as well as the frequent modal change of dynamic verbs in the imperfective (cf. 2.3.4.) supports the proposed narrative status of the respective clause.

The SF construction on the other hand requires conjunction *lē* which cannot be related to the narrative conjunction as such, but is rather segmentally identical with the NP coordinating conjunction *lè* 'and, with'. This structural similarity among the two "*le*-type" junctors and the very strict "double subject" constraint in SF might be an indication for a semantico-syntactically closer conceptualization of this construction as one single information structural unit compared to the evidently biclausal NSF organization with *tè*.

3.5 Dagbani

Like in Buli, the grammaticalization of the NC clause is restricted to NSF in Dagbani. The so-called "focus marker" *kà* is in fact just a conjunction at the beginning of a NC clause which requires verb tone paradigm C (for dynamic verbs). In NSF, a subject constituent must always follow the morpheme *kà* while coreferent subjects in two sequential clauses via clause junction *kà* are ungrammatical (cf. 2.3.5.). Since there are no constitutive structural differences between the non-focal part in NSF construction and the basic NC clause, *kà* is in both contexts still analyzed as conjunction by us, though it has some potential for grammaticalization into FM.

Interesting is the parallel to Buli found in Dagbani, insofar as here the "emphatic" marker *N* in SF resembles the NP coordinating conjunction *nì~ǹ* 'with, and'. It seems that Dagbani has a similar tendency towards a closer and
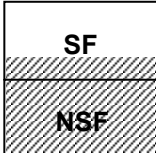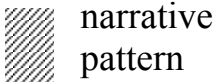
more intraclausal organization of SF compared to NSF and hence does not make use of the narrative pattern with biclausal coordination in SF.

4   Comparative Summary

As we have shown, there are striking similarities on the morphosyntactic level between the non-focal part of focus constructions and NCs, although the relevant structural parameters diverge even in our small language sample due to typological subtraits. Hence, in some of the languages – namely Akan, Buli, and Dagbani – grammatical verb tone must be taken into account in order to identify the non-focal part of focus constructions as NC. All of the languages make use of special morphological means. Apart from clausal conjunctions this also concerns suppletive pronouns. Not surprisingly it is Ewe, a language known for its pronominal specialization including logophoric forms, that provides the NC and the focus construction based on it with more than just one "dependent" pronominal form.

A typologically interesting picture in our small language sample is displayed by the distribution of the narrative structures as such in *ex-situ* focus constructions, as shown in table 2.

Table 2

On the one hand, in the Kwa languages studied, including Lelemi, the narrative pattern is more or less overtly extending into SF constructions. In the two Gur languages studied here on the other hand, SF constructions do not participate in the narrative pattern. Schwarz (in preparation) shows that in languages of this group SF rather tend to be represented by a syntactically more hierarchical (head

internal) relative construction and that the distribution of the two *ex-situ* focus constructions can be accounted for on discourse organizational grounds taking the notion of topic into account.

Having concentrated here on a deeper insight into the narrative structures that have emerged in (N)SF throughout our sample, we claim that a clausal conjunction as used to coordinate sequentials in narration does also function as device to link together focus constituent and non-focal part in a non-hierarchically way. Such focus constructions are consequently to be considered as basically biclausal, even if the clausal status of the initial clause with the focused constituent is not reflected throughout. In some of the languages, the inceptive stages of grammaticalization processes of the clause-initial conjunction into FM can be perceived, a grammaticalization chain that may even stretch further into a copula-like predicative morpheme as noticed by Stassen (1997: 85). The actual stage of such grammaticalization chain in our sample languages is shown in table 3:

Table 3

	CNJ	$\rightarrow$	**FM**	$\rightarrow$	**COP**
Ewe	*éyē*		*yé, -é*		*yé*
Akan	*nà*		*(nà)*		--
Lelemi	*nà*		*(nà)*		--
Buli	*tè*		--		--
Dagbani	*kà*		*(kà)*		--

As can be seen in the table, Akan and Lelemi exhibit homophone morphemes which could be a result of borrowing from Akan to Lelemi since loans from Akan are common in the Togo mountain area.

Three of the languages, namely Akan, Lelemi and Dagbani display the same pattern insofar as they have a conjunction which has been interpreted by some authors (Boadi 1974, Ameka 1992, Olawsky 1999) as a right-adjacent FM. According to us, the respective morphemes do have the potential for such a

functional split, but that stage has probably not yet been reached, since we could not notice any relevant categorial or structural changes of the conjunction towards a FM.

As for the Buli conjunction, there are no indications at all that it might take the grammaticalization path into a focus or predicative marker in the near future. Responsible for that is first its restriction to NSF, a fact that the Buli conjunction shares with the respective Dagbani conjunction. Second, the Buli clause conjunction is in affirmative focal contexts relatively often counter-balanced by the predicative marker respectively FM *ká* left to the focus constituent, while such an affirmative counterpart is missing in Dagbani. If the focused constituent is negated, all five languages make however use of negative copula forms. We conclude that the rarer the copula forms in affirmation are, the higher are the chances for reanalysis of the clausal conjunction as FM.

Contrary to the rather inceptive stage of grammaticalization if existent at all in most of the languages, there seems to have been a longer development in Ewe. Here, the original conjunction already shows signs of erosion when functioning as FM and it is even often suffixed to the constituent in focus.

As noted in 3.3., in Lelemi the conjunction *nà* has taken a special direction in grammaticalization. Together with the high tone born by the subject prefix in other syntagmata, it has become a "relative past" tense marker in SF. Such development from a conjunction denoting the accomplishment of actions to a past marker was also shown by Hopper (1979) for Malay, an Austronesian language.

Table 4

	CNJ	→	**"Relative Past"**
Lelemi	*nà*		*ná* (← *nà* + ´)

Our aim here was to defend the existence of a steady narrative pattern in *ex-situ* focus constructions and to outline the diachronic implications of the narrative hypothesis as an alternative to cleft and movement approaches. It has been

shown that not only the linguistic expression of the in-focus part, but also that of the out-of-focus part is important for an adequate analysis of *ex-situ* focus constructions and that the functional load verb morphology including tone has in African languages in this respect should not be underestimated.

Reference

Aboh, Enoch Oladé. 2004. *The morphosyntax of complement-head sequences. Clause structure and word order patterns in Kwa:* Oxford studies in comparative syntax. Oxford, New York: Oxford University Press.

Allan, Edward Jay. 1973. *A grammar of Buem: The Lelemi language,* School of Oriental and African Studies, University of London: PhD Thesis.

Ameka, Felix. 1992. Focus constructions in Ewe und Akan: A comparative perspective. In *Proceedings of the Kwa comparative syntax workshop* M.I.T., eds. Chris Collins and Victor Manfredi, 1-15. Cambridge.

-- 2004. Grammar and cultural practices: The grammaticalization of triadic communication in West African languages. *Journal of West African Languages* 30:5-28.

Bearth, Thomas. 1993. Satztyp und Situation in einigen Sprachen Westafrikas. In *Beiträge zur afrikanischen Sprach- und Literaturwissenschaft,* ed. W.J.G. Möhlig, 91-104. Köln.

-- 2002. Fokus im Akan - Outline. Handout of a presentation at the Linguistic Colloquium of the Seminar of African Studies at the Humboldt-University, Berlin, 28.05.2002.

Bearth, Thomas, et al. 2002. Ali Akan. Version 56. CD-ROM.

Blench, Roger. 2001. Comparative Central Togo: what have we learnt since Heine? Paper presented at *32nd Annual Conference on African Linguistics,* Cambridge.

Boadi, Lawrence A. 1974. Focus-marking in Akan. *Linguistics* 140:5-57.

Clements, George N. 1977. Four tones from three: The extra-high tone in Anlo Ewe. In *Language and linguistic problems in Africa. Proceedings of the VII. conference of African linguistics,* eds. Paul F. A. Kotey and Haig Der-Houssikian, 168-181. Columbia: Hornbeam Press.

Duthie, Alan S. 1996. *Introducing Ewe linguistic patterns.* Accra: Ghana University Press.

Fiedler, Ines, and Schwarz, Anne. (to appear). Focus or Narrative Construction? In *Focus Strategies: Evidence from African Languages*, eds. Enoch Aboh, Katharina Hartmann and Malte Zimmermann. Berlin: de Gruyter.

Frajzyngier, Zygmunt. 2004. Tense and aspect as coding means for information structure: a potential areal feature. *Journal of West African Languages* 30:53-67.

Givón, Talmy. 1990. *Syntax. A functional-typological introduction.* Vol. 2. Amsterdam: John Benjamins.

Green, Melanie. 1997. Focus and copula constructions in Hausa, School of Oriental and African Studies, University of London: Ph.D.Diss.

Heine, Bernd, and Reh, Mechthild. 1984. *Grammaticalization and reanalysis in African languages.* Hamburg: Helmut Buske.

Hiraiwa, Ken. 2003. Relativization in Buli. In *Studies in Buli Grammar. Working Papers on Endangered and Less Familiar Languages*, ed. Michael Kenstowicz, 45-84: MIT Working Papers in Linguistics.

Hopper, Paul J. 1979. Some observations on the typology of focus and aspect in narrative language. *Studies in Language* 3:37-64.

Hopper, Paul J., and Traugott, Elizabeth Closs. 1993. *Grammaticalization.* Cambridge: Cambridge Univ. Press.

Labov, Wilhelm. 1972. The transformation of experience in narrative syntax. In *Language in the Inner City,* 354-96. Philadelphia: University of Philadelphia Press.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39:463-516.

Olawsky, Knut J. 1999. *Aspects of Dagbani grammar. With special emphasis on phonology and morphology*: LINCOM studies in African linguistics; 41. München, Newcastle: Lincom.

Rongier, Jacques. 1997. Langues autonomes du Togo, entre Gur et Kwa?, Ms.

Schachter, Paul. 1973. Focus and relativization. *Language* 49.

Schiffrin, Deborah. 1994. *Approaches to discourse*: Blackwell Textbooks in linguistics. Oxford UK Cambridge USA: Blackwell.

Schwarz, Anne. 2004. Aspekte der Morphosyntax und Tonologie im Buli. Mit Schwerpunkt auf dem Buli von Wiaga. Dissertation, Humboldt-Universität zu Berlin.

-- (in preparation). Topic/Focus-Intersection in discourse: Distribution of copulative and narrative patterns in Gur.

Stassen, Leon. 1997. *Intransitive predication*: Oxford studies in typology and linguistic theory. Oxford u.a.: Clarendon Press.

Westermann, Diedrich. 1930. *A study of the Ewe language.* London: Oxford University Press.

Wilson, André A. 1963. Relative constructions in Dagbani. *Journal of African Languages* 2:139-144.

Abbreviations

CNJ	conjunction
COP	copula
DEF	definite (marker)
DEM	demonstrative (pronoun)
DEP	dependent (pronoun)
DISJ	disjunctive (pronoun)
dyn	dynamic (verb)
F	focus constituent
FM	focus marker
IND	indefinite (marker)
NC	narrative clause
NSF	non-subject focus (construction)
PROG	progressive marker
PRS	present (tense)

RC   RC

REL   relative (tense)

SF   subject focus (construction)

stat   stative (verb)

TMA   tense-mood-aspect

Ines Fiedler and Anne Schwarz
Humboldt-Universität zu Berlin
SFB 632 „Informationsstruktur"
Sitz: Mohrenstr. 40-41
Unter den Linden 6
10099 Berlin
Germany
sfb632.b1@rz.hu-berlin.de
www2.hu-berlin.de/gur_und_kwa_fokus

Diskurspragmatische Faktoren für Topikalität und Verbstellung in der ahd. Tatianübersetzung (9. Jh.)[*]

Roland Hinterhölzl, Svetlana Petrova und Michael Solf

Humboldt-Universität zu Berlin

The paper presents work in progress on the interaction between information structure and word order in Old High German based on data from the Tatian translation (9th century). The examination of the position of the finite verb in correspondence with the pragmatic status of discourse referents reveals an overall tendency for verb-initial order in thetic/all-focus sentences, whereas in categorical/topic-comment sentences verb-second placement with an initial topic constituent is preferred. This conclusion provides support for the hypothesis stated in Donhauser & Hinterhölzl (2003) that the finite verb form in Early Germanic serves to distinguish the information-structural domains of Topic and Focus. Finally, the investigation sheds light on the process of language change that led to the overall spread of verb-second in main clauses of modern German.

Keywords: information structure, verb placement in Germanic, language change

1   Zielsetzung

Der vorliegende Beitrag untersucht die Rolle der Informationsstruktur bei der Herausbildung der Wortstellungsregularitäten im heutigen Deutsch. Er ist ganz wesentlich von der Notwendigkeit bestimmt, eine Forschungslücke zu schließen − eine Forschungslücke, die ausgerechnet einer der wichtigsten und am intensivsten erforschten Zentralfragen der deutschen Syntax anhaftet, der Frage

---

nach der Herausbildung und Ausbreitung der Verbstellungsregularitäten im modernen Standarddeutschen.

Eine systematische und empirisch fundierte Untersuchung der Informationsstruktur in den früheren germanischen Sprachen steht in der bisherigen sprachhistorischen Forschung noch völlig aus. Zwar findet man sowohl in den klassischen junggrammatischen Untersuchungen von Braune (1894), Wackernagel (1891/92) und Behaghel (1932) als auch in der strukturalistischen Arbeit von Fourquet (1974) verstreute Hinweise darauf, dass informationsstrukturelle Kategorien Einfluss auf die Wortstellung der ältesten germanischen Sprachstufen nehmen[1], und ganz ähnliche Überlegungen werden auch in der Arbeit von Bean (1983) angestellt, die sich konkret mit der Entwicklung der Wortstellung im Englischen befasst. Diese Beobachtungen fließen jedoch nicht in ein kohärentes Gesamtbild zusammen, das Informationsstruktur innerhalb der Grammatik verortet und in die Erklärung von Sprachwandel einbezieht. Einen ersten Versuch in diese Richtung stellt der Ansatz von Kiparsky (1995) dar, der informationsstrukturelle Komponenten in ein Modell syntaktischen Wandels integriert.

Auffällig ist jedoch, dass die genannten Arbeiten insbesondere dann auf Faktoren der Informationsstrukturierung verweisen, wenn sie sich mit dem Problem der Verbstellung befassen. Bekanntlich zeichnet sich das finite Verb in den früheren Stufen der germ. Überlieferung durch eine hohe Positionsvielfalt im Satz aus und unterscheidet sich darin ganz wesentlich von der viel restringierteren Verbsyntax in den modernen Einzelsprachen. Dabei lässt die bisherige diachrone Syntaxforschung sowohl bei der Erklärung der Situation in der früheren Überlieferung als auch in der Frage nach der Herausbildung der

---

[1] Einen ausführlichen Überblick über die bisherige Forschungssituation geben Donhauser & Hinterhölzl (2003).

heutigen Stellungseigenschaften des finiten Verbs einen klaren Zusammenhang zwischen Verbstellung und Informationsstruktur erkennen.

Eine Arbeit, die diesen Aspekt der historischen deutschen Syntax zuletzt weiter verfolgt hat, ist die Untersuchung zur linken Satzperipherie im ahd. Isidor bei Schlachter (2004). Die informationsstrukturelle Analyse eines kleinen Belegkorpus zeigt, dass die jeweils unterschiedliche Stellung des finiten Verbs (satzfinal vs. medial) in konjunktional eingeleiteten *dass*-Sätzen mit der inhaltlichen Hervorhebung von Teilen der Aussage korreliert. Schlachter zieht daraus den Schluss, dass das finite Verb in den untersuchten Sätzen des ahd. Isidors ein Mittel der informationsstrukturellen Gliederung darstellt, indem es die Fokusdomäne einer Aussage von anderen Teilen trennt, die den Hintergrund bilden.

Die Auffassung, dass das finite Verb im ältesten Deutsch als Mittel der Identifizierung informationsstrukturell relevanter Domänen fungiert und deshalb eine hohe Stellungsvarianz aufweist, wird durch Daten gestützt, die Donhauser & Hinterhölzl (2003) anderen Texten der ahd. Überlieferung stichprobenweise entnehmen. Dieser Befund verstärkt die Annahme, dass das finite Verb im früheren Deutsch Leistungen im Bereich der informationsstrukturellen Gliederung der Aussage übernimmt, d. h. an der Herstellung der Topik-Kommentar- bzw. Fokus-Hintergrund-Gliederung von Äußerungen im Diskurs beteiligt ist.

Das Ziel der vorliegenden Untersuchung ist es, diese Hypothese auf der Grundlage eines umfangreichen Belegkorpus aus der ahd. Überlieferung zu überprüfen. Eine solche Untersuchungsaufgabe erfordert zwei wesentliche methodische Vorüberlegungen. Zum einen setzt sie die Auswahl geeigneter ahd. Textzeugnisse voraus, die verlässliche Aussagen über die Interaktion von Verbstellung und Informationsstruktur an einer möglichst großen und für den

Originalzustand des Ahd. repräsentativen Datenmenge erlauben. Zum anderen muss die Methodik der informationsstrukturellen Analyse den Besonderheiten der historischen Korpusdaten angepasst werden. Da hier Informationen über die prosodische Gestaltung der Äußerungen fehlen, muss die Bestimmung informationsstruktureller Größen wie Topik und Fokus vor allem über die Einbeziehung pragmatischer Eigenschaften aus der Kontextanalyse erfolgen.

2   Methodische Vorüberlegungen

2.1   Textwahl und textspezifische Besonderheiten

Für die Zwecke der vorliegenden Untersuchung wurde der ahd. Tatian (St. Gallen Cod. 56) ausgewählt. Es handelt sich um eine Evangelienharmonie, die vom syrischen Kleriker Tatian bereits im 2. Jh. n. Ch. zusammengestellt und schon früh in andere Sprachen übersetzt worden ist. Auf der Grundlage einer lateinischen Version entstand um 830 in Fulda auch eine althochdeutsche (ostfränkische) Übersetzung, die von sechs in unterschiedlicher Intensität beteiligten Schreibern aufgezeichnet worden ist. Der heute in St. Gallen aufbewahrte Kodex überliefert den Text als Bilingue, in der althochdeutscher und lateinischer Text spaltenweise nebeneinander stehen, vgl. Faksimile-Seite (Anhang).

Ausschlaggebend für die Auswahl des ahd. Tatians sind dessen Umfang, Alter und Textsorte: Wir haben mit dem Tatian einen der umfangreichsten und ältesten Texte der ahd. Überlieferung vor uns, der darüber hinaus – im Gegensatz zu der ebenfalls umfangreichen Überlieferung des Evangelienbuchs von Otfrid von Weißenburg – in metrisch freier Form verfasst worden ist.

Eine besondere Herausforderung bei der Arbeit am ahd. Tatian stellt die Tatsache dar, dass es sich dabei um einen Übersetzungstext handelt, der

syntaktisch in hohem Maße die mitüberlieferte lateinische Vorlage nachahmt. Der hohe Grad an Originalabhängigkeit äußert sich in der wortgetreuen Wiedergabe der Vorlage über weite Textstrecken und in der Verwendung von syntaktischen Konstruktionen, die in den germanischen Sprachen nicht heimisch sind. Typisch für die lateinisch beeinflusste Lehnsyntax des ahd. Tatians ist etwa der Gebrauch des Dativus absolutus als Entsprechung zum lateinischen Ablativus absolutus, vgl. (1)

(1)   & ihesu / baptizato & orante / confestim ascendit de aqua,
        Inti themo heilante / gitoufitemo Inti b&ontemo /sliumo úfarsteig fon
        themo uuazzare. (ahd. T 48, 30–31)
        ‚Nachdem der Heiland getauft worden war, stieg er betend schnell aus
        dem Wasser'

Um die Zahl der aus einer derartigen Nähe zur Vorlage folgenden Fehlschlüsse möglichst gering zu halten, wird die Untersuchung zunächst grundsätzlich auf Differenzbelege beschränkt, auf diejenigen Stellen also, an denen die althochdeutsche Übersetzung von der Wortstellung der Vorlage signifikant abweicht. Solche Differenzbelege finden sich in der ahd. Tatian-Übersetzung erstaunlich oft. Ihr hoher Stellenwert für die Beschreibung der Tatian-Syntax wurde bereits in den frühesten Untersuchungen erkannt, vgl. Ruhfus (1897). In neuester Zeit werden Differenzbelege im ahd. Tatian in der systematischen Abhandlung von Dittmer & Dittmer (1997) gezielt zum Gegenstand der Untersuchung.

Von den zahlreichen Fällen syntaktischer Abweichungen sind für die vorliegende Untersuchung besonders solche Fälle interessant, die die relative Stellung von Konstituenten in Bezug auf das finite Verb betreffen. Dazu gehört die Änderung der im Original vorgegebenen Konstituentenabfolge ebenso wie die Hinzufügung und Auslassung von Konstituenten gegenüber dem Latein. Für

die Zwecke unserer Untersuchung weniger hilfreich sind die wesentlich zahlreicheren Umstellungen innerhalb von NPen (Umstellung von adjektivischem bzw. Genitivattribut und Nomen u. Ä.), weshalb sie vorerst aus der Datenerhebung ausgeschlossen werden. Im Allgemeinen macht die Orientierung auf Differenzbelege das Nebeneinander von fremdsprachigem Ausgangstext und Übersetzung für die Ermittlung von authentischem ahd. Material optimal nutzbar.

Eine weitere Besonderheit, die die Übersetzungstechnik des ahd. Tatians prägt und bei der Analyse seiner Syntax unbedingt Beachtung finden muss, liegt darin, dass jede Zeile des ahd. Textes ziemlich genau das Material in der entsprechenden lateinischen Zeile wiedergibt. Dieses Prinzip der zeilenweisen Übersetzung wird im Tatian konsequent beachtet und schränkt die Möglichkeit für Umstellungen der Konstituentenabfolge zum Original gravierend ein, vgl. Masser (1997 a und b). Eine neue, handschriftennahe Ausgabe des Textes, die Masser (1994) vorgelegt hat, bietet im Vergleich zu älteren Ausgaben die Möglichkeit, diesem wichtigen Übersetzungsprinzip bei der Behandlung syntaktischer Phänomene im ahd. Tatian Rechnung zu tragen.

2.2   Methodik der informationsstrukturellen Analyse

Ein grundlegendes Problem für die informationsstrukturelle Analyse an historischen Korpora stellt das weitgehende Fehlen direkter prosodischer Information dar. Deshalb müssen insbesondere die pragmatischen Eigenschaften der Konstituenten berücksichtigt werden, die für die Identifikation informationsstruktureller Grundparameter wie Topik und Fokus in Betracht kommen. Dazu gehört prototypisch der pragmatische Status von Diskursreferenten. In Hinblick darauf darf die Bekanntheit bzw. Vorerwähntheit von Information als Voraussetzung für Topikalität bzw. Topikfähigkeit von

Konstituenten gelten, andererseits steht die Neuheit bzw. Nicht-Vorerwähntheit von Satzmaterial mit informationsstruktureller Relevanz in Verbindung und gehört daher in die Domäne des Präsentationsfokus ('Neuinformationsfokus') des Satzes.

Als Diskursreferenten werden in Anlehnung an Karttunen (1976) Individuen (Personen, Events, Objekte, Fakten etc.) aufgefasst, auf die in einem fortlaufenden Diskurs mit einem koreferenten Pronomen oder einer koreferenten definiten NP zurückverwiesen werden kann. Dies ist dann der Fall, wenn impliziert wird, dass die bezeichnete Entität in der textinternen Welt existiert und damit als Bezugspunkt in einem zusammenhängenden Diskurs zur Verfügung steht.

Über die Bestimmung des pragmatischen Status von Diskursreferenten liegt in der Literatur eine umfangreiche Diskussion vor, die eine über die Dichotomie von *gegeben* vs. *neu* hinausgehende Klassifikation begründet, vgl. Prince (1981). Unsere Untersuchung übernimmt das in Dik (^{2}1997, 315 ff.) vorgeschlagene Schema, das ausgehend vom informationellen Status der Diskursreferenten im jeweiligen Kontext folgende pragmatische Funktionen unterscheidet, die mit der Zuordnung informationsstruktureller Größen wie Topik und Fokus eng zusammenhängen:

- Ersteinführung von Diskursreferenten
- Anaphorische Referenz
- Nichtvorerwähnte, jedoch kontextuell erschließbare Referenten
- Wiederaufnahme von Diskursreferenten

Die implementierte Arbeitsmethodik sieht als ersten Schritt die Ermittlung des pragmatischen Status der Diskursreferenten aus der Kontextanalyse vor. Danach werden Beobachtungen über die Verwendung von Mitteln, die den Diskursstatus von Referenten markieren, angestellt. Vor allem interessiert dabei, ob in

Abhängigkeit vom pragmatischen Status von Diskursreferenten ein bestimmter Verbstellungstyp bevorzugt auftritt (V/1, V/2, V/end). Schließlich soll bei der Systematisierung der Beobachtungen über die Korrelation zwischen Diskursstatus und Verbstellung auch der Einfluss der Textstruktur auf die Wortstellung betrachtet werden.

3   Verbstellung und pragmatischer Status von Diskursreferenten im ahd. Tatian

3.1   Ersteinführung von Diskursreferenten

Die Untersuchung wendet sich zunächst Sätzen zu, in denen eine Entität erstmalig als Diskursreferent etabliert wird. Eine gesonderte Klasse von Sätzen, die ausschließlich auf die Funktion der Ersteinführung von Diskursreferenten spezialisiert ist, liegt im Fall der sog. Existentialkonstruktionen und Präsentationssätze (presentational sentences) vor. Diese zeichnen sich dadurch aus, dass sie typischerweise textinitial oder am Anfang eines neuen Erzählabschnitts auftreten, mit einer kleinen Gruppe nicht-agentivischer, zustandsbezeichnender Verblexeme vorkommen und dabei oft ein rahmensetzendes Adverbiale, d. h. eine Orts- oder Zeitangabe bzw. ein Expletivum (dt. *es* oder engl. *there*) in satzinitialer Position aufweisen.

Eine charakteristische Eigenschaft von Existentialkonstruktionen und Präsentationssätzen besteht darin, dass sie einen Kommentar über den in ihnen erwähnten Diskursreferenten nicht selbst abgeben, sondern ein solches im darauffolgenden Diskurs erst möglich machen. Prädikationsstrukturell betrachtet handelt es sich dabei also um Sätze ohne ein explizites Topik (Lambrecht 1994, 137–146, 177–181) bzw. um Sätze, in denen die Durchführung der Topik-Kommentar-Gliederung vollständig unterbleibt (Drubig 1992). Durch seine

Neuheit im Diskurs ist der erwähnte Referent Bestandteil der Domäne des Neuinformationsfokus, die in diesem Fall den ganzen Satz umfasst (all-new bzw. all-focus-Sätze). Präsentationssätze haben demnach – wie textinitiale Sätze im Allgemeinen auch – eine thetische Lesart.

Im ahd. Tatian zeichnet sich in Existentialkonstruktionen und Präsentationssätzen eine deutliche Präferenz für die Anfangsstellung des finiten Verbs ab. Die Konstituente, die den einzuführenden Diskursreferenten bezeichnet, folgt in diesem Fall dem finiten Verb nach.

In einem Teil der Belege liegt die Anfangsstellung des Verbs bei Ersteinführung von Diskursreferenten bereits im Lateinischen vor. Bemerkenswert ist jedoch, dass das Ahd. eine klare Tendenz dazu aufweist, die im Latein vorliegende V/1-Struktur bei Ersteinführung von Diskursreferenten aufrechtzuerhalten, vgl. (2)–(3):

(2)   **Fuit** in diebus herodis regis / iudeę <u>quidam sacerdos</u> / […]/ & <u>uxor illi</u>
        uuar [sic!] In tagun herodes thes cuninges / Iudeno <u>sumer biscof</u> / […]/
        Inti <u>quena Imo</u> (ahd. T 25, 29 ff.)
        ‚Es war zu der Herrschaftszeit von König Herodes von Judäa ein Priester
        […] und dessen Frau'

(3)   **Erat** autem <u>quidam languens</u> / lazarus a b&hania
        uuas <u>sum siocher</u> / lazarus fon b&haniu (ahd. T 228, 27 f.)
        ‚Es war ein kranker Mann, Lazarus von Bethanien.'

Noch stärker fällt jedoch ins Gewicht, dass die Anfangsstellung des Verbs bei Ersteinführung von Diskursreferenten auch in Differenzbelegen, d. h. nur in den ahd. Sätzen und gegen die Wortstellung der Vorlage, vorzufinden ist. In (4) und (5) wird eine mediale Verbstellung der lat. Vorlage zugunsten einer V/1-Stellung im Ahd. verändert, die neu einzuführenden Diskursreferenten *hirta*

‚Hirten' und *uuitua* ‚eine Witwe' erscheinen in den ahd. Belegen gegen das Original nach dem finiten Verb:

(4)   Et <u>pastores</u> **erant** In regione eadem
        uuarun thô <u>hirta</u> In thero lantskeffi[2] (ahd T 35, 29)
        ‚Da waren Hirten in jener Gegend'

(5)   <u>Vidua</u> autem quaedam **erat** / In ciuitate illa
        uuas thar ouh sum <u>uuitua</u> / In thero burgi (ahd. T 201, 2)
        ‚Es war dort auch eine Witwe in dieser Stadt'

Die Struktur solcher Sätze kann durch die in (6) angegebene Formel repräsentiert werden:

(6)   Ersteinführung von Diskursreferenten in Präsentationssätzen

        V/1: $_{FOC}[Vfin \ldots DR_{neu} \ldots]$

Neben den Fällen von verbinitialer Stellung bei Ersteinführung von Diskursreferenten finden sich in derselben Kontextfunktion auch V/2-Sätze, die links vor dem finiten Verb ein rahmensetzendes Adverbiale aufweisen, vgl. die Zeit- bzw. Ortsadverbiale ahd. *thô* ‚da, damals' in (7) und *thar* ‚dort' in (8):

(7)   & ecce <u>homo</u> **erat** In hierusalem.'
        senonu tho **uuas** <u>man</u> In hierusalem.' (ahd. T 37, 23)
        ‚Und siehe, da war ein Mann in Jerusalem'

(8)   **erant** autem ibi <u>lapideę hydrię</u>
        thar **uuarun** <u>steininu uuazzarfaz</u> (ahd. T 81, 26)
        ‚Dort waren steinerne Wasserfässer'

---

[2] Das finite Verb wird fett angegeben, der Diskursreferent, der Gegenstand der Analyse ist, wird unterstrichen.

Fälle der Ersteinführung von Diskursreferenten, die in V/2-Strukturen mit vorangestellten Frame-Adverbialen realisiert werden, können unter der in (9) angegebenen Formel systematisiert werden:

(9) Ersteinführung von Diskursreferenten

Frame + V/2 *tho/thar* $_{FOC}$[Vfin … DR$_{neu}$ ...]

Beim Vergleich von (6) und (9) wird deutlich, dass die Beispiele von V/2-Stellungen mit den in (2)−(5) aufgezeigten V/1-Sätzen vor allem eines gemeinsam haben: Die Konstituente, die dem neu eingeführten Diskursreferenten entspricht, erscheint in beiden Strukturtypen gleichermaßen rechts vom finiten Verb. Die postverbale Stellung von Ausdrücken, die neue Diskursreferenten etablieren, wird in V/2-Strukturen des Ahd. auch gegen das Latein erzeugt. Das wird in Fällen wie (7) besonders deutlich, wo die Stellung von ahd. *man* ‚ein Mann' dem finiten Verb gegen das Original nachfolgt.

Für eine Dominanz der V/1-Stellung über die *thô/thar*-V/2-Stellung bei Ersteinführung von Diskursreferenten spricht die Tatsache, dass sich Belege finden lassen, in denen ein Zeit- bzw. Ortsadverbiale gegen die Originalvorlage eingefügt wird, und zwar nach dem finiten Verb, vgl. (10)–(11):

(10) **& erat** <u>anna proph&issa</u>
     uuas thô thâr <u>anna uuizzaga</u> (ahd T 38, 22)
     ‚Es war damals dort die Prophetin Anna'

(11) **& erat** <u>quidam regulus</u>
     uuas thar <u>súm rihtari</u> (ahd. T 90, 10)
     ‚Es war dort ein Richter'

Dieselbe postverbale Stellung neuer Diskursreferenten begegnet ferner in Sätzen, in denen sie als Objekte transitiver Verben realisiert werden, vgl. (12):

(12)   Et praeteriens **uidit** <u>hominem</u> cecum / a nauitate [sic!]
        furfarentj **gisah** <u>man</u> blintan / fon giburtj (ahd. T 220, 13 f.)
        ‚Vorbeigehend sah er einen Mann, der von Geburt an blind war.'

Schließlich lassen sich bei Ersteinführung von Diskursreferenten im ahd. Tatian gelegentlich auch weitere Abfolgemuster beobachten; diese können jedoch als Ausnahmen erklärt werden. Sie sind entweder durch die Einhaltung des Zeilenprinzips wie in (13) bedingt oder stellen Nachbildungen des Originals dar, dazu (14):

(13)   ecce <u>defunctus</u> / **efferebatur**
        senu <u>arstorbaner</u> / **uúas** gitragan (ahd. T 84, 22 f.)
        ‚Siehe, es wurde ein Toter getragen'

(14)   <u>multae uiduae</u> **erant** / in diebus heliae in israhel
        <u>manago uuituuuvn</u> **uuarun**/ in heliases tagon in israhel (ahd. T 114, 27f.)
        ‚Es waren viele Witwen in den Tagen Elias' in Israel'

3.2   Anaphorische Referenz

In diesem Abschnitt soll die Situation in Sätzen betrachtet werden, in denen eine Aussage über bereits etablierte Diskursreferenten getroffen wird. Hier unterscheiden wir zwischen Fällen der linearen thematischen Progression bei einem singulären Diskursreferenten (lineare thematische Progression) und dem sog. Topikwechsel, d. h. dem selektiven oder kontrastiven Wechsel zwischen mehreren bereits eingeführten Diskursreferenten.

3.2.1 Lineare thematische Progression

In dieser Kategorie sind Fälle versammelt, in denen ein bereits eingeführter Diskursreferent zum Gegenstand einer unmittelbar nachfolgenden Aussage wird. In den Begrifflichkeiten der Prager Schule liegt hier der klassische Fall der linearen thematischen Progression vor: Das Rhema des vorangehenden Satzes wird zum Thema des Folgesatzes. Die Weiterführung des Diskursreferenten erfolgt durch pronominale Wiederaufnahme oder durch Verwendung einer koreferenten definiten NP.

Anders als bei der Ersteinführung von Diskursreferenten, bei der eine klare Präferenz für die V/1-Stellung sowie für die postverbale Realisierung der jeweiligen Diskursreferenten beobachtet wurde, sind bei der unmittelbaren Weiterführung bereits bekannter Referenten bei einem Großteil der Differenzbelege V/2-Strukturen zu finden. Dabei wird die Stelle links vor dem finiten Verb durch den bereits vorerwähnten Diskursreferenten selbst besetzt. Diese Konstellation ist in Minimalpaaren von aufeinanderfolgenden Sätzen zu beobachten, in denen der erste Satz eine V/1-Stellung bei Ersteinführung eines Diskursreferenten aufweist und der zweite Satz denselben Diskursreferenten in die satzinitiale Stelle eines V/2-Satzes übernimmt, vgl. (15):

(15)   **Fuit** […] quidam sacerdos / […] / & uxor illi […] / **erant** autem iusti
        ambo ante deum
        uuar [sic!] […] sumer biscof /[…] / Inti quena Imo […] / siu **uuarun**
        rehtiu beidu fora gote (ahd. T 25, 29–26, 3)
        ‚Es war […] ein Bischof […] und dessen eine Frau […]. Sie waren beide
        rechtschaffen vor Gott.'

Die Prädikationsstruktur dieser Sätze erlaubt es, sie als kategorische Sätze mit einer Topik-Kommentar-Gliederung zu interpretieren. Die Konstituente, die im Anschlusssatz links vom finiten Verb steht, ist ihrem pragmatischen Status nach

Topik im Sinne der familiarity- und aboutness-Konzepte (vgl. die Zusammenstellung in Frey 2000: 137 f.). Das finite Verb eröffnet die Domäne, in der ein Kommentar zum Satztopik erfolgt, und markiert gleichzeitig den Beginn der Domäne, die neue Information im Diskurs präsentiert.

Die Struktur, die sich damit für Sätze mit anaphorischer Referenz im Ahd. ergibt, kann folgendermaßen dargestellt werden:

(16)   Anaphorische Referenz

$$\text{V/2} \quad _{\text{TOP=BGR}}[\text{DR}_{\text{giv}}] \; _{\text{FOC=COMMENT}}[\text{Vfin} \dots]$$

Dieses Abfolgemuster in kategorischen Sätzen mit bekannten Diskursreferenten wird im Ahd. auf vielfältige Weise gegen das Original konstruiert. Zunächst werden Belege angeführt, in denen die Übersetzung zwar die im Original vorliegenden Konstituenten übernimmt, ihre Abfolge jedoch konsequent in Richtung der in (16) ausgewiesenen Struktur umstellt:

(17)   ego sum pastor bonus. bonus pastor / animam suam **dat** pro ouibus suis
         ih bin guot hirti. guot hirti / **tuot** sina sela furi siniu scaph.
         (ahd. T 225, 16 f.)
         ‚Ich bin ein guter Hirte. Der gute Hirte gibt seine Seele für seine Schafe.'

(18)   ecce defunctus / efferebatur. filius unicus / matris suae. & haec uidua
         erat.
         senu arstorbaner / uúas gitragan einag sun / sinero muoter Inti thiu **uuas**
         uuituuua (ahd. T 84, 22 ff.)
         ‚Siehe, da wurde ein Toter getragen, der einzige Sohn seiner Mutter, und
         diese war Witwe'

In diesen Belegen wird die Spät- bzw. Endstellung des Originals zugunsten einer V/2-Stellung im Ahd. verändert, wobei die einzige Stelle vor dem finiten Verb der unmittelbar vorerwähnten Konstituente als Topik der Aussage vorbehalten

bleibt. Der Kommentar darüber bzw. die Domäne der Neuinformation wird mit dem finiten Verb eingeleitet.

Ferner wird im Ahd. die in (16) angegebene V/2-Struktur bei anaphorischer Referenz auch durch Einsetzung von im Original fehlenden Konstituenten erzeugt. Dazu gehört die Hinzufügung der finiten Kopula in Zweitstellung bei der Wiedergabe elliptischer Kopulativkonstruktionen des Lateins oder bei Auflösung synthetischer Verbformen wie in (19) bzw. die Einsetzung des Subjektpronomens links vom finiten Verb wie in (20):

(19)   lazarus **Infirmabatur**
          lazarus **uuard** cumig (ahd. T 229, 3)
          ‚Lazarus ist krank geworden'

(20)   & uxor tua […] / pari& tibi <u>filium</u>. / […] / & **erit** tibi gaudium &
          exultatio […] / **erit** enim magnus coram domino
          Inti thin quena […] / gibirit thir <u>sun</u>. / […] / Inti <u>her</u> **ist** thir gifeho Inti
          blidida […]/ <u>her</u> **ist** uuârlihho mihhil fora truhtine (ahd. T 26, 25 ff.)
          ‚Und deine Frau […] wird dir einen Sohn gebären. […] Er wird dir
          Freude und Wonne sein. [...] Er wird wahrlich groß vor Gott sein'

Bei einem Großteil der ermittelten Belege kommen die Umstellung und Hinzufügung von Konstituenten gegen das Latein kombiniert zur Anwendung, vgl. (21):

(21)   [Auf die Frage hin: „Wie sind dir deine Augen wieder aufgetan
          worden?":]
          <u>Ille homo</u> qui dicitur / ihesus. lutum **fecit** & unxit / oculos meos
          <u>ther man</u> thiedar ist giquetan / heilant. <u>her</u> **t&a** leimon inti salbota / minu
          ougun (ahd. T 221, 9 f.)
          ‚Dieser Mann, der da genannt wird Heiland, er rührte Schlamm ein und
          salbte [damit] meine Augen'

Starke Evidenz dafür, dass die in (16) aufgezeigte V/2-Struktur mit der pragmatischen Funktion der anaphorischen Referenz zusammenhängt, liefern Differenzbelege in eingeleiteten Nebensatzstrukturen des Ahd. Unmittelbar vorerwähnte Konstituenten, die im Original postverbal erscheinen, werden im Ahd. systematisch in die Position links vor dem finiten Verb gestellt:

(22)   & non erat <u>illis</u> filius. eo quod / **ess&** <u>elisab&h</u> sterilis
         Inti niuuard <u>In</u> sun. bithiu uuanta / <u>elisab&</u> **uuas** unberenti (T 26, 6 f.)
         ‚Und sie hatten keinen Sohn, weil Elisabeth unfruchtbar war'

(23)   Lucerna corporis. est <u>oculus</u>. / si **fuerit** <u>oculus tuus</u> simplex. / totum
         corpus tuum lucidum **erit**.
         liohtfaz thes lihhamen ist <u>ouga</u> / oba <u>thin ouga</u> **uuirdit** luttar / thanne **ist**
         al thin lihhamo liohter (ahd. T 69, 21 ff.)
         ‚Das Licht des Körpers ist das Auge. Wenn dein Auge hell wird, dann
         wird auch dein ganzer Körper licht sein'

Dieselbe Art der Konstituentenumstellung wird in eingeleiteten Nebensätzen des Ahd. selbst dann vorgenommen, wenn der entsprechende lat. Satz die V/end-Stellung aufweist, vgl. (24)−(25):

(24)   [...] ueni / ut […] <u>qui uident</u> caeci **fiant**
         [...] quam ih […] / thaz [...] <u>thie dar gisehent</u> daz <u>sie</u> **sin** blinte
         (ahd. T 224,4 ff.)
         ‚Ich kam […] damit diejenigen, die sehen, blind werden'

(25)   domine ego credidi / quia tu es christus filius dei / <u>qui</u> In mundum
         uenisti
         trohtin ih giloubta. / thaz thu bist crist gotes sun. / <u>thie dar</u> **quam** In
         mittilgart (T 231, 18 ff.)
         ‚Herr, ich glaubte, dass du Christus, der Sohn Gottes bist, der auf die Welt
          kam = und dieser kam auf die Welt'

Sowohl vor der Maßgabe der maximalen Anlehnung an die Originalwortfolge in diesem Text als auch vor dem Hintergrund der Annahme, dass die V/end-

Stellung im eingeleiteten Nebensatz des Ahd. – so wie im heutigen Deutsch auch – die kanonische Verbstellung repräsentiert (Lenerz 1984), ist die Evidenz in (24)−(25) überaus erstaunlich. Für die wissenschaftliche Auswertung dieses Textes sowie für die verfolgte Argumentationslinie sind diese Belege in zweierlei Hinsicht besonders von Belang. Zum einen relativieren sie die von Dittmer & Dittmer (1998, 18) aufgestellte Behauptung, dass der ahd. Tatian bezüglich seiner Vorlage an keiner Stelle in Richtung einer für das heutige Deutsch untypischen Syntax abweiche. Zum zweiten stehen sie pragmatisch und strukturell vollkommen in Einklang mit den anderen vorab präsentierten Fällen einer Weiterführung bekannter Diskursreferenten: Es ist auffällig, dass die Konstituente links vom finiten Verb auch hier vorerwähnt ist, während die Domäne nach dem finiten Verb das inhaltliche Hauptgewicht der Mitteilung trägt.

Diese Daten bestätigen die Beobachtung, dass V/2-Stellungen im Ahd. regelmäßig mit der pragmatischen Funktion der anaphorischen Referenz korrelieren, und stützen weitgehend die Hypothese, dass die Verbstellung im Ahd. als Mittel der informationsstrukturellen Gliederung der Aussage eingesetzt wird.

3.2.2 Topikwechsel

Oft bezieht sich der Satz nicht auf einen einzigen, sondern auf mehrere Diskursreferenten, die vorerwähnt und daher im Gedächtnis der Diskurspartizipanten aktiviert, d. h. salient sind. Es stehen damit mehrere potentiell geeignete Topik-Kandidaten im Diskurs bereit. Von entscheidender Bedeutung ist dabei die Tatsache, dass diesen Diskursreferenten gemäß ihrer syntaktischen Realisierung im vorangehenden Diskurs ein unterschiedlicher Grad an Salienz zukommt, der für die Gestaltung des Nachfolgesatzes von Bedeutung ist. Ein gängiges Modell, die Salienz von Diskursreferenten auf einer

Skala abzubilden, liegt in der Abstufung Subjekt > direktes Objekt > indirektes Objekt > Rest vor. Dabei nimmt man an, dass der in der vorangehenden Äußerung salienteste Diskursreferent seinen Status auch in der nachfolgenden Struktur behält, wenn ein Wechsel in der Salienzskala nicht durch besondere Mittel angezeigt wird. Zu diesen Mitteln zählen etwa die Genuskongruenz, die Wahl von Anaphern mit einem höheren Grad an sprachlicher Explizitheit (Ø-Anapher, Personalpronomen, volle NP als Eigenname, nominaler Ausdruck etc.) bzw. der Unterschied zwischen Personal- und Demonstrativpronomen im Deutschen, vgl. Lambrecht (1994, 204 f.). Auf diese Weise ist dafür gesorgt, dass der jeweilige anaphorische Ausdruck mit dem richtigen Antezedens identifiziert und die Äußerung angemessen interpretiert werden kann.

Die vorliegenden Beobachtungen konzentrieren sich auf Strategien zur Kennzeichnung eines Wechsels in der Salienzskala von anaphorischen Ausdrücken, wobei grundsätzlich zwischen einem selektiven Wechsel von Diskursreferenten und einem zusätzlich mit Merkmalen der kontrastiven Gegenüberstellung verknüpften Wechsel unterschieden wird.

(+ contrast / + selective)

Bei Ausdrücken mit diesen Merkmalen handelt es sich um kontrastive Topiks. In einem Teil der Belege wird diese pragmatische Funktion mit der für die anaphorische Referenz typischen V/2-Stellung auch gegen die Wortstellung in der Vorlage realisiert. Der Kontrast wird möglicherweise auch durch prosodische Hervorhebung der in Kontrast gesetzten Konstituenten erzielt, wofür sich in einem historischen Korpus jedoch kaum sichere Angaben ermitteln lassen. In (26) wird allerdings eine Kontrastbetonung der Konstituenten ziemlich sicher durch die Setzung der deiktischen Personalpronomen lat. *tu* ‚du' und lat. *nos* ‚wir' in Subjektstellung nahegelegt.

(26)   <u>tu</u> discipulus illius **sis** / <u>Nos</u> autem moysi discipuli **sumus**
         <u>thu</u> **sís** sín iungiro / <u>uuir</u> **birumes** moyseses iungiron (ahd. T 223, 7 f.)
         ‚Du mögest sein Jünger sein, wir sind Jünger des Moses'

Während in (26) der eigentliche Kontrastmarker – die Partikel lat. *autem* ‚aber' – nicht mitübersetzt wurde, findet sich in anderen Belegen mit kontrastiver Lesart eine entsprechende Markierung durch Partikel bzw. Adverbien.

Eine interessante Beobachtungsmöglichkeit bietet die ahd. Übersetzung der kontrastiven Partikel lat. *autem* in der häufig vorkommenden Wendung lat. *ego autem dico* ‚ich aber sage'. Die Wiedergabe von *autem* durch *thanne* im Ahd. hängt in diesen Fällen ziemlich regelmäßig mit dem Vorliegen einer kontextuell deutlich erschließbaren Kontrastivität zusammen. Fest steht, dass der ahd. Text das Adverb *thanne* als Entsprechung zu lat. *autem* stets dort enthält, wo sich der Sprecher (Jesus) in irgendeiner Weise inhaltlich von den zitierten Aussagen der früheren Propheten distanziert. Das betrifft etwa die Auffassung vom Schwören in (27), von der persönlichen Rache in (28) und von der Scheidung in (29):

(27)   Dictum est autem.' Quicumque dimiserit. / uxorem suam.' d& illi
         libellum repudii; / <u>Ego autem</u> **dico** uobis., / quia omnis qui dimiserit /
         uxorem suam […] / facit eam moechari.
         íz ist giq&an só uuér so fúrlaze. / sina quenun. gebe iru buoh thanatribes.
         / <u>thanne ih</u> **quidu** íu / thaz thero giuuelih thie furlazit / sina quenun […] /
         tuot sia furligan (ahd. T 64, 6 ff.)
         ‚Es ist auch gesagt: „Wer sich von seiner Frau scheidet, der soll ihr einen
         Scheidebrief geben." Ich aber sage euch: „Wer sich von seiner Frau
         scheidet, […], der macht, dass sie die Ehe bricht" […]'

(28)   Audistis quia dictum est antiquis. / non perierabis […] <u>ego autem</u>
         dico uobis., / non iurare omnino
         Ir gihórtunt thaz then alton giq&an uúas / nifúrsuueri thih. […] <u>thanne</u>
         <u>ih</u> **quidu** íu / thaz mán zi thuruhslahti nisuuere (ahd. T 64, 13 ff.)
         ‚Ihr hörtet, dass zu den Alten gesagt worden ist: „Du sollst nicht falsch
         schwören!" Ich aber sage euch, dass man überhaupt nicht
         schwören soll.'

(29)   Audistis quia dictum est / oculum pro oculo. [...] <u>ego autem</u> **dico** uobis.,
         / non resistere malo
         Ir gihórtut thaz giqu&an ist / ouga furi ouga. [...] / <u>thanne ih</u> **quidu** íu
         / thaz ír niuuidarstant& ubile. (ahd. T 64, 29 ff.)
         ‚Ihr hörtet, dass gesagt worden ist: „Auge um Auge." Ich aber sage euch,
         dass ihr dem Übel nicht widerstreben sollt.'

Bei inhaltlicher Übereinstimmung mit der Aussage der Propheten, so etwa
bezüglich der Auffassung von Totschlag (30) und Ehebruch (31), bleibt lat.
autem im Ahd. unübersetzt:

(30)   Audistis quia dictum est. / antiquis.' non occides. […] / <u>ego autem</u> **dico**
         uobis. / quia omnis qui irascitur / fratri suo. reus erit iudicio
         ír gihórtut thaz giqu&an uúas. / then alton. niuúis manslago / [...] <u>ih</u>
         quidu íu / thaz iogiuuelih ther sih gibilgit / zi sinemo bruoder. ther ist
         sculdig duomes (ahd. T 62, 21 ff.)
         ‚Ihr habt gehört, dass zu den Alten gesagt wurde: „Du sollst nicht töten."
         Ich sage euch, dass, wer seinem Bruder zürnt, des Gerichts schuldig ist.'

(31)   Audistis quia dictum est / antiquis.' non moechaberis., / <u>Ego autem</u> **dico**
         uobis., / quoniam omnis qui uiderit mulierem / ad concupiscendum eam.'
         / Iam moechatus est eam in corde suo.
         Ir gihortut thaz giqu&an ist / then alton nifurligi thíh / <u>ih</u> **quidu** íu / thaz
         iogiuuelih thiethar gisihit uúib / sie zigeronne / iu habet sia forlegana in
         sinemo herzen (ahd. T 63, 18 ff.)
         ‚Ihr hörtet, dass zu den Alten gesagt wurde: „Du sollst nicht
         ehebrechen!" Ich sage euch, dass wer eine Frau ansieht, um sie zu
         begehren, mit ihr schon die Ehe gebrochen hat in seinem Herzen'

Der syntaktische Status dieser Kontrastmarkierung ahd. *thanne* ist zweifelhaft. Sie kann sowohl als Modifizierer innerhalb der NP als auch als selbständige syntaktische Konstituente mit Satzgliedstatus aufgefasst werden. Deshalb lässt sich auch die Verbstellung in solchen Fällen (V/2 vs. V3) nicht eindeutig bestimmen.

(- contrast; + selective)

Selektiver Topikwechsel wird im ahd. Tatian ebenfalls durch Partikeln oder Adverbien signalisiert, wobei zugleich auch eine Tendenz zur Früherstellung des finiten Verbs zu beobachten ist, vgl. (32) und (33). Der V/2-Status der Sätze ist allerdings weitgehend unklar, vgl. besonders (33):

(32)   (Nach dem Erscheinen des Engels, während Zacharias das Rauchopfer bringt:)
&zacharias turbatus **est**
thanan tho <u>zacharias</u> **uuard** gitruobit (ahd. T 26, 20)
‚Zacharias aber erschrak davor'

(33)   **mansit** autem <u>maria</u> cum illa / quasi mensibus tribus & reuersa **est** / In domum suam, / <u>Elisab&h</u> autem Impl&um est / tempus pariendi
uuon&a <u>maria</u> mit Iru / nah thri manoda Inti **uuarb** / zi Ira hûs, / <u>Elisab&h</u> uuârlihho **uuard** gifullit / zît ziberanne (ahd. T 30, 13 ff.)
‚Maria blieb bei ihr [Elisabeth] etwa drei Monate lang und ging zurück in ihr Haus. Für Elisabeth aber erfüllte sich die Zeit zu gebären'

Von einer Signalfunktion der Partikeln und Adverbien bei einer Verschiebung der Salienz von Diskursreferenten kann auch in Fällen häufiger lexikalischer Wiederholungen ausgegangen werden, die die Herstellung der anaphorischen Beziehung gefährden, vgl. (34):

(34)   Quis uestrum hab& <u>amicum</u> / & ibit ad <u>illum</u> media nocte / & dicit <u>illi</u>.
,<u>Amice</u>.' / commoda mihi tres panes / quoniam amicus meus uenit de uia
/ ad mé [...] <u>& ille</u> de intus **dicat** […]
Vuelih íuuer hab& <u>friunt</u> / Inti ferit zi <u>Imo</u> In mittero naht / Inti quidit
<u>imo</u>. <u>friunt</u> / Intlih mír thriu brót / uuanta mín friunt quam fon uúege zi
mir. […] <u>her</u> <u>thanne</u> fon innana **quede** […] (ahd T 72, 13 ff.)
,Wer$_i$ von euch hat einen Freund$_j$, und geht zu ihm$_j$ mitten in der Nacht
und sagt ihm$_j$: „Freund$_j$, gib mir$_i$ drei Brote, denn mein Freund$_k$ ist von
der Reise zu mir$_i$ gekommen [...]" und er$_j$ würde dann von drinnen sagen:
[…]'

3.3   Nichtvorerwähnte, jedoch erschließbare Diskursreferenten

In diesem Teil betrachten wir Sätze mit Diskursreferenten, die im
vorangehenden Kontext zwar nicht explizit erwähnt sind, aber in einer
Analogie- oder Teil-Menge-Beziehung zu einem anderen, vorab eingeführten
Diskursreferenten stehen. Solche Diskursreferenten sind kontextuell bzw. aus
dem Weltwissen des Sprechers inferierbar und stellen geeignete Topik-
Kandidaten bereit, über die im jeweiligen Kontextbezug eine Aussage erwartbar
ist, vgl. (vgl. Prince 1981; Dik [2] 1997, 323 f.).

In diesen Fällen ist im ahd. Tatian – wie bei den Fällen der anaphorischen
Referenz – eine Präferenz für V/2-Stellungen zu beobachten, wobei inferentiell
erschließbare Diskursreferenten in der für kontextuell vorerwähnte Topiks
typischen Stelle unmittelbar vor dem finiten Verb vorkommen:

(35)   & <u>nomen eius</u> elisab&h
Inti <u>ira namo</u> **uuas** elisab&h (ahd. T 26,2)
,Und ihr Name war Elisabeth'

In (35) ist die Konstituente ahd. *ira namo* ,ihr Name' auf die vorab eingeführte
Person ahd. *quena* ,[seine] Frau' bezogen. Die Nennung des Namens gilt nach
der Einführung der Handlungspersonen einer Erzählung durchaus als erwartbar.

Das gegen das Latein eingesetzte finite Kopulaverb steht hier zwischen der durch Inferenzbeziehung topikal zu deutenden Konstituente und der neuen Information im Prädikatsnomen und fungiert somit als trennendes Glied zwischen dem Topik und dem Neuinformationsfokus der Aussage. Solche Vorkommen fallen informationsstrukturell und syntaktisch mit den Fällen der Weiterführung bekannter Diskursreferenten zusammen und weisen dabei die Struktur von (16) auf.

3.4 Wiederaufnahme von Diskursreferenten

In diesem Teil werden Sätze betrachtet, in denen bekannte Diskursreferenten nach einer gewissen Unterbrechung in den Diskurs wiederaufgenommen werden.

Aufgrund des Umstands, dass sich Eigenschaften der Wiederaktivierung mit Merkmalen anderer pragmatischer Funktionen überschneiden (Ersteinführung, Wechsel von Diskursreferenten), ist es durchaus problematisch, der Wiederaufnahme den Status einer eigenständigen pragmatischen Klasse zuzuweisen. Darum überrascht auch nicht, dass die ahd. Sätze, in denen die Rede nach einer Unterbrechung erneut auf einen vorab erwähnten Diskursreferenten kommt, keine einheitliche syntaktische Struktur erkennen lassen. Zum einen finden sich hier Fälle von V/1-Stellungen in Korrespondenz-(36) aber auch in Differenzbelegen (37), wobei der wiedereingeführte Referent spät im Satz, rechts vom finiten Verb erscheint:

(36)   **erat** autem & <u>iohannes</u> baptizans
        Vuas ouh tho <u>iohannes</u> toufenti (ahd. T 56, 23)
        ‚Auch Johannes taufte dort'

(37)   & ecce <u>angelus domini</u>
         quam thara <u>gotes engil</u> (ahd. T 35, 32)
         ‚Da kam Gottes Engel'

In anderen Fällen, in denen die Wiedereinführung von Diskursreferenten begegnet, treten allerdings auch V/2-Sätze auf, wobei der wiedereingeführte Diskursreferent die Stellung vor dem finiten Verb einnimmt:

(38)   <u>puer autem</u> **crescebat** & confortabatur / spiritu
         <u>ther kneht</u> **uuvohs** Inti uuard gistrengisot / geiste (T 32, 6 f.)
         ‚Dieser Knabe wuchs und wurde gestärkt im Geiste'

Hier wird durch die Hinzufügung des einfachen Demonstrativpronomens ahd. *ther* ‚dieser' gegen das Latein die Bekanntheit bzw. Identifizierbarkeit des Diskursreferenten suggeriert, was eine Topik-Lesart der entsprechenden Konstituente nahe legt.

       Darüber hinaus werden Partikeln bzw. Adverbien verwendet, die ähnlich wie im Fall des Wechsels aktiver Diskursreferenten eine Verschiebung auf der Salienzskala signalisieren; syntaktisch ist die Verbstellung nicht eindeutig als V/2 zu bestimmen:

(39)   <u>Puer autem</u> **crescebat**
         <u>ther kneht</u> uuârlihho **uuvohs** (ahd. T 42, 8)
         ‚Dieser Knabe wuchs wahrlich'

3.5   Zwischenbilanz

Die vorangehende Untersuchung hat gezeigt, dass sich bestimmte Verbstellungstypen im Tatian in der Regel mit bestimmten pragmatischen Funktionen von Diskursreferenten in Verbindung bringen lassen. So findet man bei der Ersteinführung von Diskursreferenten, etwa in Präsentationssätzen und Existentialkonstruktionen, vorzugsweise die V/1-Stellung. Dagegen begegnet

die V/2-Stellung regelmäßig in Sätzen mit kontextuell vorerwähnten bzw. erschließbaren Referenten, die in der Position vor dem finiten Verb stehen. Damit bestätigt sich am Beispiel des Tatians die Annnahme, dass zwischen der Stellung des finiten Verbs und dem pragmatischem Status von Diskursreferenten ein Zusammenhang besteht.

Wenn man dieses Fazit auf die Ebene der Informationsstruktur der betreffenden Sätze überträgt, lassen sich die Aussagen über die Verteilung der Verbstellungsmuster in Abhängigkeit vom informationellen Status von Diskursreferenten weiter präzisieren. Die Existenzialkonstruktionen und Präsentationssätze, die sich im Ahd. durch eine hohe Präferenz für V/1-Abfolgen auszeichnen, gelten als all-focus-Sätze ohne Satztopik. V/2-Stellungen sind wiederum typisch für Sätze mit Diskursreferenten, die sowohl im Sinne des familiarity- als auch im Sinne des aboutness-Konzepts als Satztopiks identifizierbar sind. In diesem Fall nimmt der bekannte Diskursreferent, d. h. der Topikausdruck, die Stellung unmittelbar vor dem finiten Verb ein.

Damit steht auf der Basis einer breiteren empirischen Erhebung fest, dass V/1- und V/2-Stellungen im Ahd. primär eine funktionale Opposition bilden, die im Rahmen der informationsstrukturellen Gliederung der Äußerung angesiedelt ist. Die Früherstellung des finiten Verbs im Ahd. gegenüber der V/end-Stellung, die als basisgeneriert angenommen wird, lässt sich mit Leistungen im Bereich der informationstrukurellen Gliederung der Aussage identifizieren: Genauer gesagt besetzt das finite Verb im Ahd. sowohl in V/1- als auch in V/2-Sätzen dieselbe Position, nämlich die Position, die den Beginnn der Domäne des Neuinformationsfokus auszeichnet. Der entscheidende Unterschied zwischen den Verbstellungstypen besteht darin, dass V/1-Strukturen all-focus-Sätze ohne ein explizites Topik darstellen, während V/2-Sätze mit satzinitialer Topik-Konstituente als in Topik und Kommentar bzw. Fokus und Hintergrund

gegliedert erscheinen, wobei das finite Verb hier die Domänen von Topik und Kommentar/Fokus voneinander abgrenzt.

Den ausgewiesenen Früherstellungen – V/1 und V/2 im Ahd. – kann die folgende gemeinsame Struktur zugewiesen werden:

(40)   a.   V/1 $_{FOC}[V_{fin}...DR_{new}...]$

       b.   V/2 $_{TOP=BGR}[DR_{giv/acc}]$ $_{FOC=COMMENT}[V_{fin}...]$

Der Wandel in der Verbsyntax vom Ahd. zum Nhd. hin besteht demnach in der Generalisierung des V/2-Musters über den Bereich von Sätzen mit einem expliziten referentiellen Satztopik hinaus sowie in der damit einhergehenden Marginalisierung von V/1-Strukturen auf textinitiale Sätze in metrischer Dichtung und umgangssprachlicher Verwendung.

4   Gegenprobe

Die Zwischenbilanz aus der empirischen Untersuchung der Zusammenhänge zwischen Verbstellung und pragmatischem Status von Diskursreferenten im ahd. Tatian, die eine systematische Präferenz der V/1- und V/2-Verbstellung in unterschiedlichen Kontexten ergab, soll an dieser Stelle einer Gegenprobe unterzogen werden. Zum einen sollen Fälle diskutiert werden, die vorerwähnte und daher allgemein topikfähige Konstituenten enthalten, dennoch aber eine V/1- bzw. *thô*+V/2-Struktur aufweisen. Zum anderen sollen regelmäßige Vorkommen von V/2-Stellungen im Ahd. erörtert werden, in denen die einzige Konstituente vor dem finiten Verb nicht vorerwähnt ist und damit nicht zwingend als Topik im Sinne des aboutness-Konzepts einzuordnen ist. Dazu gehören direkte Fragesätze mit satzinitialem Fragewort und Sätze mit einer nicht-vorerwähnten, daher fokalen Konstituente vor dem finiten Verb.

4.1 Weitere Verwendungsbereiche von V/1 im ahd. Tatian

Auffällig ist, dass die V/1-Stellung im ahd. Tatian über die regelmäßige Verwendung in der Funktion der Ersteinführung von Diskursreferenten hinaus in zahlreichen weiteren Kontexten vorkommt. Interessanterweise enthalten diese Gruppen von Sätzen bereits vorerwähnte Diskursreferenten, welche jedoch nicht in der Position vor dem finiten Verb, sondern rechts davon erscheinen. Dieser Befund scheint das von uns gewonnene Fazit über die Verteilung der Verbstellung in Abhängigkeit von der pragmatischen Funktion der Sätze zunächst zu widerlegen und soll deshalb hier genauer untersucht werden.

Starke Verbreitung findet das V/1-Muster in Sätzen mit Fortbewegungsverben. Ein Teil von ihnen dient in der Tat zur Ersteinführung von Diskursreferenten, vgl. (41):

(41)   **uenit** <u>mulier</u> / de samaria haurire aquam;
        quam tho <u>uuib</u>$_{new}$ / fon samariu sceffen uuazzar (ahd. T 130, 30 f.)
        ‚Kam da eine Frau aus Samaria, um Wasser zu schöpfen'

Die textpragmatische Funktion dieser Äußerung stimmt demnach weitgehend mit der von Existentialkonstruktionen und Präsentationssätzen überein. Das trifft jedoch nicht auf alle Sätze mit Fortbewegungsverben zu. Die V/1-Stellung kommt nämlich auch dann vor, wenn der Diskursreferent in einem Satz mit einem Fortbewegungsverb bereits vorerwähnt ist, vgl. (42)–(43):

(42)   **E**t regressus **est** <u>ihesus</u>
        inti **uuidarfuor** tho <u>ther   heilant</u>$_{giv}$ (ahd. T 53, 14)
        ‚Und der Heiland kehrte da zurück'

(43)   & reuersus **est** <u>centurio</u> in domum suam
        uuarb tho <u>ther centenari</u>$_{giv}$ in sin hús (ahd. T 84, 8)
        ‚Ging da der Zenturio in sein Haus zurück.'

Ferner findet sich die V/1-Stellung im ahd. Tatian regelmäßig in Sätzen, die den Beginn eines neuen Erzählabschnitts signalisieren. Besonders häufig sind darunter die Fälle mit der Floskel ahd. *uuard tho* ‚es geschah, es trug sich zu, etc.' für lat. *factum est.* Wie (44) zeigt, tendiert das Ahd. auch dann zu einer satzinitialen Stellung des finiten Verbs, wenn das unpersönliche Verb *uuard* gemeinsam mit einem Partizip Perfekt auftritt und damit auch eine genaue Nachbildung des Originals möglich gewesen wäre:

(44)   Factum **est** autem in allio sabbato
        uuas thó giuuortan in anderemo sambaztag (ahd. T 106, 6)
        ‚Es war da geworden an einem anderen Sabbatstag'

Auch außerhalb dieser Anfangsfloskel begegnet die V/1-Stellung in Kontexten, in denen ein persönlicher Ausdruck den Beginn eines neuen Erzählabschnitts bezeichnet. Oft entspricht das dem Übergang zu einer neuen Bibelstelle und ist anhand der Konkordanzangaben leicht zu verfolgen. Zusätzlich ist der Beginn neuer Textabschnitte im Original graphisch durch Initialen in Capitalis gekennzeichnet. Obwohl in solchen Fällen die Anfangsstellung des Verbs bereits im Lateinischen dominiert, vgl. (45), lassen sich auch Differenzbelege ermitteln, bei denen am Beginn eines neuen Textabschnitts eine V/1-Abfolge im Ahd. hergestellt wird, vgl. (45)–(47):

(45)   **Rogauit** autem illum quidam / phariseus
        bat inan$_{giv}$ sum / phariseus$_{new}$ (T 126, 1–2)
        ‚Es bat ihn ein Pharisäer'

(46)   Non de omnibus u̲o̲b̲i̲s̲ **dico**
        ni **quad** i̲h̲$_{giv}$ fon í̲u̲$_{giv}$ allen (T 271, 21)
        ‚Nicht sprach ich von euch allen'

(47)   **Acceperunt** autem corpus ihesu
         Intfiengung <u>sie</u>_{giv} tho thes heilantes lichamon (T 321, 29)
         ‚Empfingen sie da des Heilands Leichnam'

Allgemein erstreckt sich die V/1-Stellung auch auf Fälle, die den Beginn neuer Situationen innerhalb desselben Textabschnitts bezeichnen, so bei inchoativer Lesart wie in (48):

(48)   <u>Phariseus</u> autem **coepit** intra se / reputans dicere
         bigonda <u>ther phariseus</u>_{giv} innan imo / ahtonti queden **(**ahd. T 126, 5 f.)
         ‚Es fing der Pharisäer an, bei sich zu sprechen'

Damit lässt sich feststellen, dass die Position des finiten Verbs im Ahd. nicht nur Leistungen im Bereich der informationsstrukturellen Gliederung der Äußerung übernimmt, sondern global als textstrukturierendes Signal eingesetzt wird, das einen Situationswechsel im Diskurs markiert.

Mit der Diskurssemantik des Situationswechsels sind weitere regelmäßige Fälle von V/1 im Tatian verbunden, die den Eintritt eines neuen physischen, psychischen oder kognitiven Zustands an bekannten Diskursreferenten bezeichnen:

(49)   factus **est** timor super <u>omnes uicinos eorum</u>
         uuard thô forhta ubar <u>alle Iro nahiston</u>_{giv} (ahd. T 31, 2)
         ‚Furcht überkam da alle ihre Verwandten'

(50)   & **sensit** corpore / quod sanata ess& a plaga
         furstuont <u>siu</u>_{giv} thó in ira lihhamen / thaz siu heil uuas fon theru suhti
         (ahd. T 95, 14 f.)
         ‚Da merkte sie an ihrem Körper, dass sie von der Krankheit geheilt worden war.

(51)   **Cognouit** ergo <u>pater</u>
         furstuont tho <u>ther fater</u>_{giv} (ahd. T 91, 2)
         ‚Da verstand der Vater'

Schließlich ist bei den V/1-Verwendungen im ahd. Tatian eine große Gruppe an Belegen, darunter auch Differenzbelegen, zu erwähnen, die verba dicendi enthalten und einen Sprecherwechsel im Dialog ankündigen. Auch hier folgen vorerwähnte Diskursreferenten den finiten Verb nach:

(52)   & respondens <u>angelus</u> / dixit ei.
         antlingota tho <u>ther engil</u>_{giv} / quad Iru. (ahd. T 28, 26 f.)
         ‚Da antwortete der Engel [und] sagte zu ihr'

(53)   & respondens <u>mater eius</u> & dixit
         antlingota thô <u>sîn muoter</u>_{giv} Inti quad (ahd. T 30, 24)
         ‚Da antwortete seine Mutter und sagte'

Die Verberststellung in diesen Fällen kann damit erklärt werden, dass eine Redeeinleitung an sich mit dem Wechsel in der generellen Beschaffenheit der Erzählsituation verbunden ist und diese Kontexte dem oben beschriebenen Situationswechsel mit V/1-Stellung funktional nahe kommen.

        Es ist offenkundig, dass das V/1-Muster im Ahd. eine ganze Reihe von Verwendungen besitzt, die in den Bereich der Textorganisation und Diskursstrukturierung führen. Die Funktion der Ersteinführung von Diskursreferenten, die oben beobachtet wurde, ist nur eine davon. Mit dieser teilen die im Folgenden aufgeführten weiteren Gruppen von V/1-Belegen auch die funktionsidentische Doppelung durch *thô*+V/2:

(54)   **adducunt** eum / ad pharisęos
         tho **leittun** sie thén / zi then pharisein **(**ahd. T 221, 15)
         ‚Da führten sie diesen zu den Pharisäern'

(55)   & repl&i **sunt** omnes / in sinagoga ira
        thó **uuvrdun** sie gifullte alle / in theru samanungu gibuluhti
        (ahd. T 115, 7)
        ‚Da wurden sie alle in der Versammlung von Zorn erfüllt'

(56)   **Dixit** autem maria
        thô **quad** maria (ahd. T 128, 18)
        ‚Da sagte Maria'

Die Sichtung des Materials legt nahe, dass bei den verschiedenen Schreibern eine unterschiedlich ausgeprägte Präferenz für *thô*+V/2 vs. V/1 in den Kontexttypen vorliegt, die einen Situationswechsel signalisieren. Eine erste Quantifizierung liegt für den Bereich der Redeeinleitung vor. Diese zeigt eine 100%ige Präferenz für *thô*+V/2 in der Redeeinleitung bei Schreiber ε; der Vergleich der Vorkommen von V/1 vs. *thô*+V/2 in derselben Textmenge bei anderen Schreibern ergab ein Verhältnis 16:3 bei α, 3:9 bei β und 1:12 bei ζ. In Bezug auf diese Fragestellung sind jedoch weitere Nachforschungen erforderlich.

4.2   V/2 in Sätzen mit präverbaler Neuinformation

Die Herausbildung der V/2-Stellung im frühesten Deutsch hängt nach unseren Beobachtungen primär mit dem Bereich der anaphorischen Referenz zusammen, wobei das finite Verb ein Satztopik topologisch vom inhaltlichen Kernstück der Aussage, der Domäne des Neuinformationsfokus trennt. Dieser Befund soll mit anderen Verwendungen der V/2-Stellung im Ahd. verglichen werden, bei denen die Konstituente vor dem finiten Verb nicht topikal ist.

Zum einen handelt sich dabei um direkte Ergänzungsfragen, die gemäß einer detaillierten empirischen Eigenuntersuchung (Petrova & Solf in Vorbereitung) bereits im ältesten Deutsch über eine sehr rigide V/2-Syntax verfügen. Diese zeigt sich darin, dass das Grundmuster bei direkten

Ergänzungsfragen, nämlich die Abfolge Fragewort-V/2-Rest, im Tatian in 202 der insgesamt 230 ermittelten Belege vorliegt, und in 125 dieser Fälle in Differenzbelegen gegen das Original erzeugt wird, vgl. (57) :

(57)   unde hoc **sciam**   wh>DO>Vfin
        uanan **uueiz** ih thaz (ahd T 27, 10)   wh>Vfin>SU>DO
        ‚Woher weiß ich das?'

Entscheidungsfragen, die statt der V/2-Struktur mit satzinitialem Fragewort ein davon abweichendes Muster aufweisen, sind bis auf 4 Differenzbelege von insgesamt 28 Vorkommen stets als Nachbildungen des Originals zu erklären und darüber hinaus vor allem auf einen einzigen Schreiber, nämlich γ, beschränkt.

Daraus folgt, dass sich im Bereich der direkten Ergänzungsfragen offenbar seit frühester Zeit eine feste V/2-Syntax ausgebildet hat. Das Problem dieses Fazits für die vorliegende Arbeit besteht darin, dass die präverbale Konstituente in direkten Ergänzungsfragen, d. h. das Fragewort selbst, nicht mit pragmatischen und informationellen Merkmalen von Topikalität in Verbindung gebracht werden kann. Wir finden m. a. W. in gewissen Bereichen des Ahd. Evidenz für eine fest ausgebildete V/2-Syntax, die mit den empirischen Beobachtungen aus der informationsstrukturellen Analyse anderer früher Vorkommen von V/2 in Deklarativsätzen nicht vereinbar ist.

Probleme scheinen ferner solche V/2-Strukturen in Deklarativsätzen zu bereiten, in denen die präverbale Konstituente neue Information trägt und daher nicht als topikal, sondern als fokal zu deuten ist. Dies ist etwa in dem folgendem Differenzbeleg der Fall, wo wir links vom finiten Verb die Domäne des engen XP-Fokus finden, rechts davon die topikale Konstituente ahd. *her* ‚er':

(58)   mortuus **est**
         toot **ist** $\underline{\text{her}}_{giv}$ (ahd. T 149, 14)
         ‚Tot ist er'

Im Vergleich wird allerdings deutlich, dass die Domäne des Neuinformationsfokus im Ahd. in einer verhältnismäßig größeren Anzahl an Differenzbelegen rechts vom finiten Verb realisiert wird, vgl. (63)–(66):

(59)   quinque enim uiros **habuisti**
         $\underline{\text{thu}}_{giv}$ **habetos** $\underline{\text{finf gomman}}_{new}$ (ahd T 132, 6)
         ‚Du hattest fünf Männer'

(60)   alii autem nequaquam. / sed similis **est** eius
         andere quadun nist / úzouh $\underline{\text{hér}}_{giv}$ **ist** imo gilih (ahd. T 221, 5–6)
         ‚Andere sagten: „Er ist es nicht, aber er ist ihm gleich."'

(61) $\underline{\text{lutum}}$ **fecit**
         $\underline{\text{her}}_{giv}$ **t&a** $\underline{\text{leimon}}_{new}$ (ahd. T 221, 10)
         ‚Er rührte Schlamm ein'

(62) $\underline{\text{demonium}}$ **hab&**
         $\underline{\text{er}}_{giv}$ **hab&** $\underline{\text{diuual}}_{new}$ (ahd. T 226, 18)
         ‚Er ist vom Teufel besessen'

Schließlich ist ein kleiner Bereich von V/2-Strukturen in Differenzbelegen des ahd. Tatians anzuführen, bei dem sich vor dem finiten Verb ein kataphorisches Element befindet, das auf einen postverbalen weiten Fokus verweist (Fokusvorwegnahme):

(63)   & hoc uobis signum, Inueni&is / Infantem pannis Inuolutum.' / &
         positum in presepio,
         thaz **sî** îu zi zeichane. thaz ir find& / kind mit tuohon biuuvntanaz.' / Inti
         gilegitaz in crippa; (ahd. T 36, 8 ff.)
         ‚Das sei euch zum Zeichen, dass ihr ein Kind finden werdet, das in
         Tücher gewickelt und in eine Krippe gelegt ist'

Die Katapher ahd. *thaz* ‚das' ist Gegenstand einer weiterführenden Erläuterung. Vor diesem Hintergrund kann argumentiert werden, dass das kataphorische Element informationsstrukturell betrachtet Merkmale eines aboutness-Topiks aufweist, sofern es eine Entität etabliert, über die eine Aussage getroffen wird.

5   Ausblick und weitere Untersuchungen

Das Fazit der empirischen Untersuchung über die Korrelation von Verbstellung und pragmatischem Status von Diskursreferenten lautete, dass die V/2-Stellung ein konstitutives Merkmal von Sätzen mit bekannten Diskursreferenten darstellt und in funktionaler Opposition zu V/1-Sätzen steht, in denen Diskursreferenten erstmals eingeführt werden. Die Gegenprobe zur Informationsstruktur bei davon abweichenden V/1- und V/2-Vorkommen in Differenzbelegen des ahd. Tatians erfordert an dieser Stelle einige Präzisierungen.

Zunächst ist für die Fälle von V/1 davon auszugehen, dass dieses Verbstellungsmuster eine breite Verwendung im Ahd. fand und in einer größeren Anzahl an Kontexten bevorzugt erscheint, die ganz allgemein einen Situationswechsel im Diskurs signalisieren. Das V/2-Muster dagegen ist auch in Differenzbelegen nachzuweisen, bei denen das finite Verb keine referentielle bzw. keine topikale Vorfeldkonstituente vom Neuinformationsfokus der Aussage trennt. Dazu zählen die Frame-Adverbiale *thô, thar* sowie die kataphorischen Elemente, die auf nachfolgende Neuinformation verweisen. Sie zeugen davon, dass bereits in ahd. Zeit ein Wandel im Gang ist, der sich in der Ausbreitung des V/2-Musters über die Domäne der anaphorischen Referenz hinaus äußert. Als ein Zeichen der fortschreitenden Ausbreitung von V/2 sind auch die festgestellten Fälle von engen XP-Foki links vor dem finiten Verb zu

deuten, die allerdings bei den Differenzbelegen zahlenmäßig deutlich hinter Fällen von postverbalem Fokus zurückstehen.

Bedeutsamer für das Verständnis der Generalisierung von V/2 ist die Evidenz bei den selbständigen Fragesätzen mit Fragewort. Für ihre Erklärung kommen aus unserer jetzigen Sicht zwei Möglichkeiten in Betracht. Die erste setzt eine Modifizierung der These von der Funktion des finiten Verbs als Mittel der Trennung der informationsstrukturellen Domänen von Topik und Fokus in Satz voraus. Eine mögliche Anpassung dieser These zielt auf die Generalisierung der Funktionsweise des finiten Verbs im Ahd. Demnach bewirkt das finite Verb die Abgrenzung einer pragmatisch konsistenten Vorfelddomäne, die entweder topikal oder fokal, jedoch nicht pragmatisch gemischt sein kann. Eine ähnliche informationsstrukturelle Differenzierungsfunktion von Adverbien im Mittelfeld, die eine Domänenkonsistenz bei der Serialisierung von Argumenten schaffen, beobachtet Musan (2002) im heutigen Deutsch. Wir können in Analogie dazu die Stellung des finiten Verbs im Ahd. als ein Mittel erklären, informationsstrukturelle Domänen nach Möglichkeit als zusammenhängende Einheiten voneinander zu trennen.

Die zweite Möglichkeit, die Generalisierung von V/2 mit der Evidenz im Bereich der direkten Fragesätze in Einklang zu bringen, geht mit der Vermutung einher, dass die Entwicklung der V/2-Stellung bei den direkten Fragesätzen und bei den Deklarativsätzen das Ergebnis zweier voneinander unabhängiger Prozesse darstellt: eines syntaktisch bedingten Prozesses bei den Fragesätzen und eines informationsstrukturell bedingten Prozesses bei den Deklarativa mit V/2-Abfolge in kategorischer Lesart. Der unterschiedliche pragmatische Status der präverbalen Elemente in diesen Bereichen fester V/2-Setzung führt zu einer Umdeutung der Vorfeldposition bei den Deklarativa. Diese wird von einer

informationsstrukturell vorbestimmten zu einer informationsstrukturell indifferenten Domäne, wobei nunmehr allein die Anzahl, jedoch nicht die Pragmatik der sie besetzenden Konstituenten von Belang ist.

Wie die Daten zeigen, ist der Generalisierungsprozess von V/2 in dem untersuchten Zeitraum bereits in Gang. Das erklärt die informationelle Vielfalt der präverbalen Elemente bei V/2-Abfolgen in Differenzbelegen des ahd. Tatians. Gleichzeitig sind jedoch klare quantitative Präferenzen zugunsten referentieller Topikelemente im Vorfeld von V/2-Strukturen zu erkennen. Dies deutet sowohl auf den funktionalen Ursprungsbereich von V/2 bei den Deklarativa mit kategorischer Lesart hin, als auch auf den Wandelprozess, der sich in der kontinuierlichen Ausbreitung des V/2-Musters auf weitere Kontextbereiche äußert.

6 Anhang
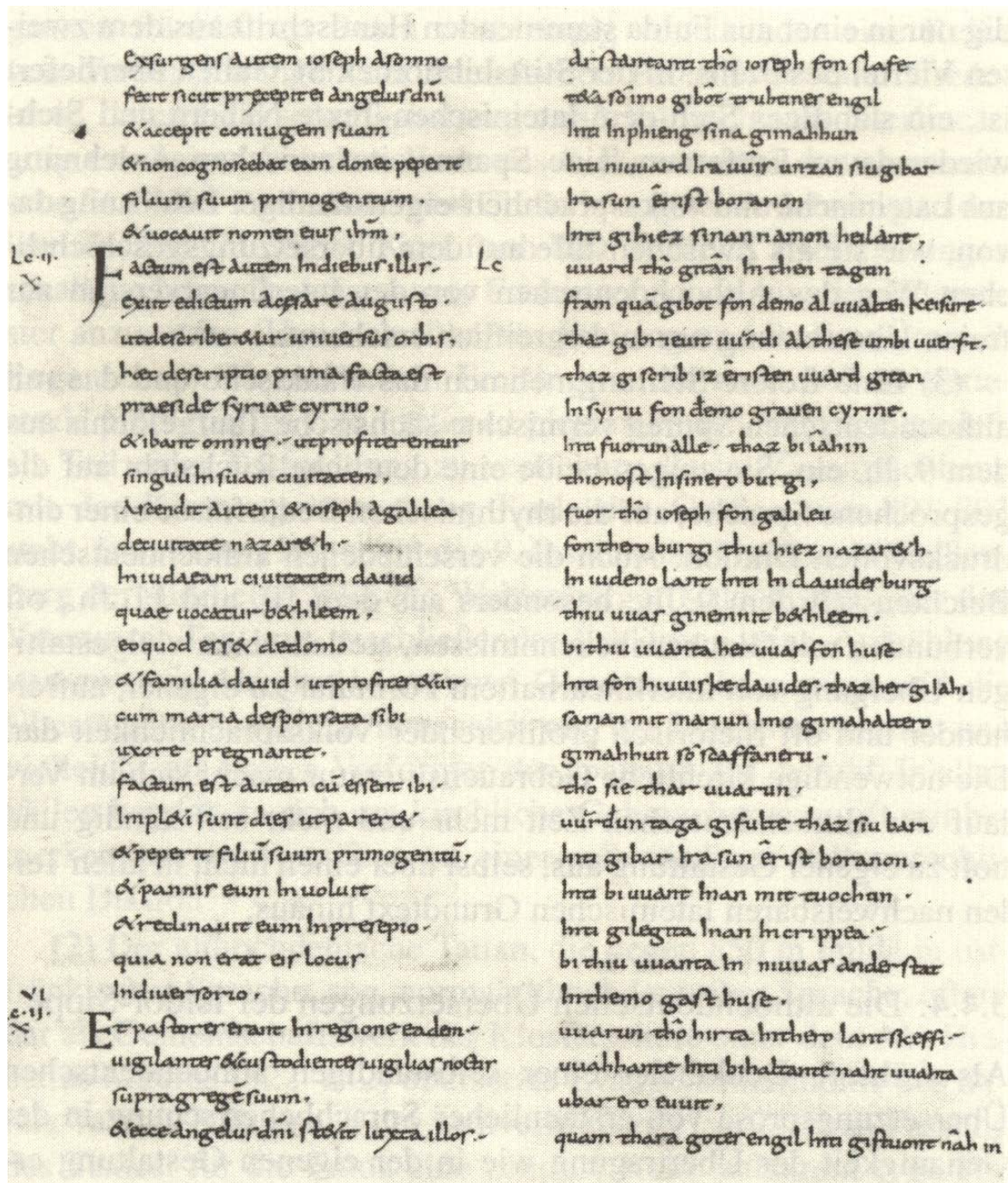
Reference

Textquellen

[T]   Die lateinisch-althochdeutsche Tatianbilingue Stiftsbibliothek St. Gallen
        Cod. 56. Hg. v. A. Masser. Göttingen: Vandenhoeck & Ruprecht, 1994.
        (= Studien zum Althochdeutschen, Bd. 25)

Wissenschaftliche Literatur

Bean, Marian C. 1983. The Development of Word Order Patterns in Old
        English. London and Canberra: Croom Helm. (=Croom Helm Linguistics
        Series)

Behaghel, Otto. 1932. Deutsche Syntax. Eine geschichtliche Darstellung. Band
        IV. Wortstellung. Periodenbau. Heidelberg: Winter. (= Germanische
        Bibliothek. I. Sammlung germanischer Elementar- und Handbücher. 1.
        Reihe: Grammatiken. 10. Bd.)

Braune, Wilhelm. 1894. Zur Lehre von der deutschen Wortstellung. In:
        Forschungen zur deutschen Philologie. Festgabe für Rudolf Hildebrand
        zum 13. März 1894. Leipzig: Veit & Co.

Dik, Simon C. 1997. The Theory of Functional Grammar. Part I: The Structure
        of the Clause. Second, revised Edition. Edited by K. Hengeveld. Berlin /
        New York: Mouton de Gruyter. (= Functional Grammar Series 20)

Dittmer, Arne und Ernst Dittmer. 1998. Studien zur Wortstellung.
        Satzgliedstellung in der althochdeutschen Tatianübersetzung. Göttingen:
        Vandenhoeck & Ruprecht. (= Studien zum Althochdeutschen, Bd. 34)

Donhauser, Karin und Roland Hinterhölzl. 2003. Die Rolle der
        Informationsstruktur bei der Herausbildung von Wortstellungsregularitäten
        im Germanischen. In: C. Féry et al. (Hgg.): Informationsstruktur. Die
        sprachlichen Mittel der Gliederung von Äußerung, Satz und Text.
        Finanzierungsantrag 2003–2007. Potsdam und Berlin: Universität Potsdam,
        173–193.

Drubig Bernhard. 1992. Zur Frage der grammatischen Repräsentation thetischer
        und kategorischer Sätze. In: J. Jacobs (Hg.): Informationsstruktur und
        Grammatik. Linguistische Berichte. Sonderheft 4, 142–195.

Fourquet, Jean. 1974. Genetische Betrachtungen über den deutschen Satzbau. In: Besch, Werner et al. (Hgg.) Studien zur deutschen Literatur und Sprache des Mittelalters. Festschrift für Hugo Moser zum 65. Geburtstag. Berlin: E. Schmidt, 314–323.

Frey, Werner. 2000. Über die syntaktische Position des Satztopiks im Deutschen. In: ZAS Papers in Linguistics 20, 137–172.

Karttunen, Laurie. 1976. Discourse referents. In: J. McCawley (Hg.): Syntax and Semantics, Vol. 7. New York: Academic Press, 363–385.

Kiparsky, Paul. 1995. Indoeuropean Origins of Germanic Syntax. In: A. Battye and I. Roberts (eds.): Clause Structure and Language Change. Oxford: Oxford University Press, 140–169.

Klein, Wolfgang und Christiane von Stutterheim. 1992. Textstruktur und referentielle Bedeutung. In: Zeitschrift f. Literaturwissenschaft und Linguistik 86, 67–92.

Lambrecht, Knud. 1994. Information Structure. Topic, Focus, and the Mental Representations of Discourse Referents. Cambridge: Cambridge University Press. (= Cambridge Studies in Linguistics, 71)

Lenerz, Jürgen. 1984. Syntaktischer Wandel und Grammatiktheorie. Eine Untersuchung an Beispielen aus der Sprachgeschichet des Deutschen. Tübingen: Niemeyer.

Lühr, Rosemarie (im Druck): Der Einfluß der klassischen Sprachen auf die germanische Grammatik. In: G. Meiser, Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft in Halle/S., 17. –23. September 2000.

Masser, Achim. 1997a. Syntaxprobleme im althochdeutschen Tatian. In: Y. Desportes (Hg.), Semantik der syntaktischen Beziehungen. Akten des Pariser Kolloquiums zur Erforschung des Althochdeutschen 1994. Heidelberg: Winter, 123–140.

Masser, Achim. 1997b. Wege zu gesprochenem Althochdeutsch. In: E. Glaser und M. Schlaefer (Hgg.): Grammatica Ianua Artium. Festschrift für Rolf Bergmann zum 60. Geburtstag. Heidelberg: Winter, 49–70.

Musan, Renate. 2002. Informationsstrukturelle Dimensionen im Deutschen. Zur Variation der Wortstellung im Mittelfeld. In: ZGL 30, 2002, 198–221.

Petrova, Svetlana und Michael Solf (in Vorbereitung): Syntaktischer Wandel und Satzmodus. Beobachtungen zur Wortstellung in direkten Fragesätzen des Althochdeutschen.

Prince, Ellen F. 1981. Towards a Taxonomy of Given-New Information. In: P. Cole (Hg.), Radical Pragmatics. New York: Academic Press, 223–255.

Ruhfus, Wilhelm. 1897. Die Stellung des Verbums im althochdeutschen Tatian. Inaugural-Dissertation zur Erlangung der Doctorwürde. Dortmund: Selbstverlag.

Schlachter, Eva. 2004. Satzstruktur im Althochdeutschen. Eine Skizze zur Position des Verbs im Isidor-Traktat des 8. Jahrhunderts. In: K. Pittner et al. (Hgg.), Beiträge zu Sprache und Sprachen 4. Vorträge der Bochumer Linguistik-Tage. München: Lincom, 179–188.

Wackernagel, Jacob. 1891/92. Über ein Gesetz der indogermanischen Wortstellung. In: Indogermanische Forschungen 1, 333–346.

Roland Hinterhölzl
Humboldt-Universität zu Berlin
Institut für deutsche Sprache und Linguistik
Sitz: Schützenstr. 21
Unter den Linden 6
10099 Berlin
roland.hinterhoelzl@rz.hu-berlin.de

Svetlana Petrova and Michael Solf
Humboldt-Universität zu Berlin
SFB 632 „Informationsstruktur"
Sitz: Mohrenstr. 40-41
Unter den Linden 6
10099 Berlin
Germany
www2.hu-berlin.de/sprachgeschichte/forschung/informationsstruktur

Focus accent, word length and position as cues to L1 and L2 word recognition [*]

Anke Sennema, Ruben van de Vijver,
Susanne E. Carroll, Anne Zimmer-Stahl

Universität Potsdam

The present study examines native and nonnative perceptual processing of semantic information conveyed by prosodic prominence. Five groups of German learners of English each listened to one of 5 experimental conditions. Three conditions differed in place of focus accent in the sentence and two conditions were with spliced stimuli. The experiment condition was presented first in the learners' L1 (German) and then in a similar set in the L2 (English). The effect of the accent condition and of the length and position of the target in the sentence was evaluated in a probe recognition task. In both the L1 and L2 tasks there was no significant effect in any of the five focus conditions. Target position and target word length had an effect in the L1 task. Word length did not affect accuracy rates in the L2 task. For probe recognition in the L2, word length and the position of the target interacted with the focus condition.

Keywords: bilingual word processing, prosodic prominence

1   Introduction

Focus, as expressed through pitch, has been identified as an important aspect of language comprehension, and a number of factors influence the assignment of focus during sentence comprehension.

It has been shown that rapid and effective processing of the accent placement in an utterance contributes to efficient comprehension of meaning (Cutler & Fodor,

1979) and that there is a robust advantage for words with predicted accent (Cutler, 1976). Listeners exploit cues in the prosodic contour preceding an accent to locate possible accent points. Since accent falls on semantically crucial, i.e. focused words, listeners actively search for focus to facilitate comprehension. Drawing the listeners' attention to certain parts of the utterance through means of prosodic prominence might lead to more detailed processing of the signal and this might lead to faster word recognition (Cutler *et al.*, 1997).

This evidence from studies of native listening inspired investigations addressing nonnative processing strategies of accentual sentence structures. A consistent and similar interaction of accent and focus was reported for native processing of English and Dutch, although this effect did not show in nonnative listening (Akker & Cutler, 2003). The recognition memory of English sentences was tested in which prosody cued meaning contrasts, and memory performance based on prosodic information was shown to be generally poor (Pennington & Ellis, 2000). After participants' attention was explicitly directed to intonation, the performance improved only on sentences with contrastive focus pairs.

From this it is not yet clear, to what extent the prosodic realization of focus play a role in L2 processing. Therefore, in a comparative analysis we investigated the role of focal accent, word length and the position of the word in the sentence on word recollection in native and nonnative recognition tasks in German L2 learners of English. If it is the case that the realization of focus through pitch enhances processing and recollection than such effects occur in short and long words and independent of the position in the sentence of the target word.

2   Study

We aimed at examining the relationship between effects of focus as realized by accent for native and nonnative listening. Furthermore, we have addressed the influence of sentence position and word length on word recognition in L2.

For L1 it has been established that accented material is processed more efficiently (Cutler, 1976). It has been shown, that L2 listeners attend more readily to the beginning and end of the sentence and that learners show a preference for the sentence beginning (Barcroft & VanPatten, 1997), (Klein, 1992). We will investigate the extent to which prosodic prominence interacts with the Sentence Location Principle of VanPatten. Our hypothesis is, in contrast to VanPatten, that learners will be most sensitive to words which are prosodically marked for focus regardless of its location in the sentence. We expect, therefore, that German L2 learners of English will recognize target words better when these occur at sentence initial or final position.

It has been argued, that stressed syllables are articulated in a clearer fashion (Cutler & Norris, 1988), and that hence it would be easier to represent its phonetics and phonological characteristics. In the case of words of more than one syllable, the listener might be able to accurately represent the properties of the stressed syllable but not of unstressed syllables, making recognition of the whole word more difficult than recognition of a word consisting of one stressed syllable. All of these issues - the influence of focus as realized by accent, position in the sentence and word length - have been addressed in the present study.

2.1   Speech materials

Target words in each language consisted of names of birds, such as *Gohl* (a German target) or *scaups* (an English target). They were judged by three German learners of English for their familiarity and balanced in the lists of

target words. The items were one-syllabled or more-syllabled. For every target one filler item was constructed. Target items were controlled for word length and sentence position, the fillers were not. The items were embedded in sentences.

To study the effects of different focus positions on word recognition, prosodic variation was triggered by using wh-questions for otherwise identical sentences. A wh-question focuses a specific constituent of the sentence and the answer to that question ought to focus the same constituent (Selkirk, 1992). The following example from the test material shows the relation between pitch accent and focus (pitch accent is marked by capital letters):

(1)   a.   "AUKS are being affected by the warming of the northern seas because they prefer cold waters. "

would be the answer to the question

      b.   "Who is being affected by the warming of the northern seas?",

whereas

(2)   a.   "Auks are being affected by the warming of the northern seas because they prefer COLD WATERS."

would be the answer to the question

      b.   "Why are auks being affected by the warming of the northern seas?"

Elicited sentences were broad focus sentences (condition B1), narrow focus realised on the target (condition N1, as in example 1a), or narrow focus realised on a constituent other than the target (condition B2, as in example 2a). In the carrier sentence the target words occurred in three different positions, i.e. initial, medial and final.

To evaluate the effect of the cues in the prosodic contour surrounding an accented word, two test conditions with spliced material were constructed. In the

first spliced condition (spliced B1), the target word of a B1 sentence was spliced into the context of a N1 sentence, i.e. an unaccented target word was spliced into the N1 context of focused constituent. In the second condition (spliced B2), the target of a B2 sentence was spliced into a N1 sentence, i.e. an unaccented and unfocused target word was spliced into a sentence with narrow focus on the target. With this material 5 separate experimental conditions with identical set-up were programmed.

2.2   Speakers and Recording procedure

A female speaker of South Eastern British English recorded the English items and a female native speaker of Standard German recorded the items in German.

For each language 24 target items and 24 foils across sentence position and syllable length were recorded in three conditions, i.e. B1, N1 B2. 75 filler sentences were constructed in B1 and B2 sentence conditions, some of which had to be used twice in the experiment to make up numbers. In addition, one token of each target and foil was recorded in isolation. Also, 12 sentences plus 3 target items were recorded for a familiarization phase at the beginning of the experiment. Digital recordings were made in a sound-proof booth, using an Audiotechnica 4033a microphone (audio sampling rate 22.05 kHz).

2.3   Listeners

104 listeners participated in the experiment. They were German students or employees at the University of Potsdam and were at an intermediate to advanced level of English proficiency. The range of age was between 18 and 43 years. All participants had all started learning English after the age of 8 and none had lived in an English speaking country for more than 24 months.

A British English control group of 45 students was tested in the UK. 28 of them were recruited from University College London and 17 from Essex

University, where they were first and second year students of Linguistics.

All participants in both language groups either received points for course requirements or were paid a small sum for their participation. At the time of the experiment no reported normal or corrected hearing and normal or corrected vision.

There were 20 German subjects in each condition N1, B1, B2 and B2 spliced and 24 German subjects in condition B1 spliced. In each condition N1, B2, B1 spliced there were 10 English controls and 13 subjects in condition spliced B2.[1]

2.4   Experimental task and procedure

A closed-set word probe detection task was built for the experiment. The task was explained to the listeners and they were instructed to pay attention to the sentences in order to do well on the word probe detection task. After that they entered a brief training session in which they heard three blocks each of 4 sentences followed by a single word. Listeners were asked to press one key on a computer keyboard when they recognized the word heard in isolation as one having occurred in one of the previous 4 sentences, and another key when the word in isolation did not occur in one of the previous 4 sentences. They were instructed to make their decision as quickly as possible.

In the trial part feedback was given on the correctness of the answers but no feedback was given during the actual test and there was no further communication with the experimenters. Subjects heard the stimuli sentences only once. Two self-timed pauses were programmed within the experiment. Each experiment (German task and English task) took about 30 minutes.

---

[1]   Due to an error only two subjects were tested for condition B1 unspliced.

The experiment had two separate parts: a part with German stimuli and one with English items. Both parts consisted of 48 blocks of four sentences each. There were 24 blocks with 'no' as the correct answer and 24 blocks with 'yes' as the correct answer. The targets were equally distributed in terms of presentation order within the blocks (pos. 1-4) and the blocks were presented in randomized order. To compensate for fatigue effects a second list was created with the blocks in the reversed order of the first list. The sentences and probes were presented at a comfortable listening level via headphones. Their responses and reaction times were recorded.

To obtain an independent measure of English proficiency, the Oxford Placement Test was administered. This is a standardized test (multiple choice) divided into two sections, Listening and Grammar. The percentage scores from the sections were used to obtain a relative measure of English proficiency of the German participants. An attempt was made to stratify our learners by proficiency level on the basis of the score achieved in the Listening part of the Oxford Placement Test and to distribute learners equally across the five focus conditions.

The experimental order was as follows: Oxford Placement Test, German part of the experiment (L1 task), and one week later, the English part of the experiment (L2 task). The English control group was tested on the English part only.

2.5 Results

Timed-out responses (a response latency of more than 2500 ms) and responses with reaction times below 300ms were discarded from the analysis. The percentage of correct probe recognition in each part of the experiment (L1 task and L2 task) was calculated for the five test conditions (Table 1).

Tab. 1: Scores (% correct) per focus condition and part of experiment for
             German subjects and English controls

*N=2

German subjects	**B1**	**N1**	**B2**	**B1 spliced**	**B2 spliced**
L1 task	93,1%	93,7%	92,9%	93,2%	94,2%
L2 task	82,6%	88,9%	85,7%	87,3%	85,3%
Controls	100%*	93,5%	86,3%	89,4%	85,4%

Accuracy scores were subjected to an analysis of variance with language as the within subjects factor (L1 vs L2) and focus as the between subjects factor. There was a significant difference between L1 and L2 [$F(1,99) = 117.21$, $p < .001$], showing that German subjects performed better in their native language L1 than in L2. Whereas there was no main effect of focus, focus tended to interact with language [$F(4,99) = 2.17$, $p = .08$]. Accuracy scores for the experimental task in the L1 and the task in the L2 in the five focus conditions are shown in Figure 1.

Fig. 1: Boxplots of accuracy scores
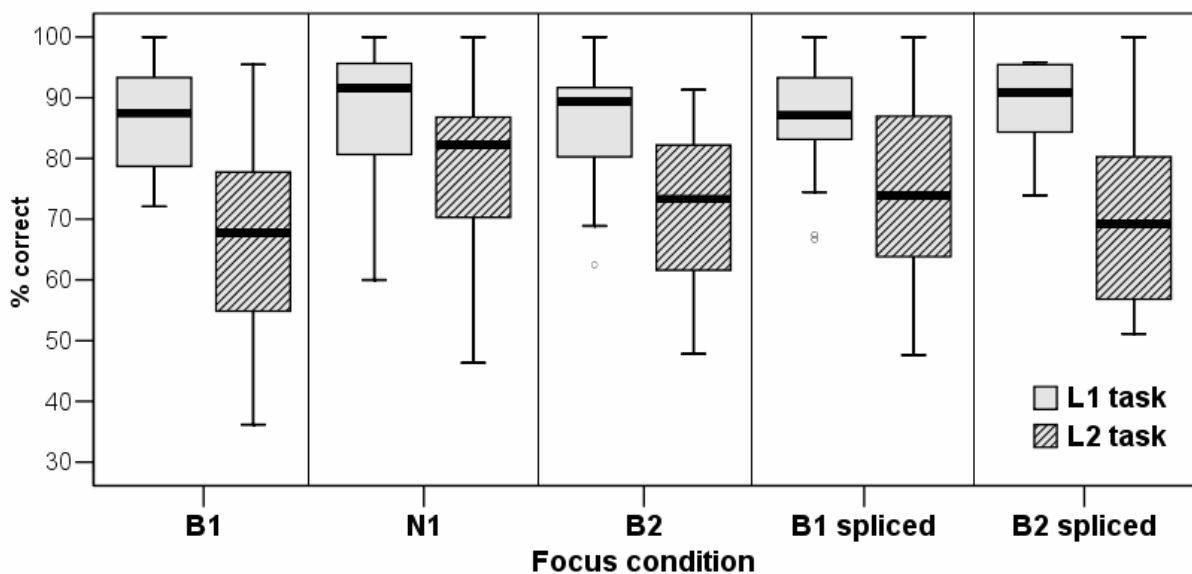
Next, the factors of word length and target position in the sentence were examined for the five conditions in each of the two experimental tasks. Accuracy scores for targets occurring in sentences were subjected to an analysis of variance with target length (mono- or polysyllabic) and target position in the sentence (initial, medial, final) as within subjects factors and focus as between subjects factor. Paired comparisons for target position examined initial and medial position against final position.

In the L1 task with broad focus (B1) there was no effect for target length, but a significant difference between the target positions [$F(2,38) = 6.032$, $p = .05$]. Paired comparisons of the final position against initial and medial position revealed a significant difference between initial and final target position [$F(1,19) = 6.032$, $p < .005$] . In the corresponding L2 task, however, there was no effect for target length and also no effect for target position (see Fig. 2).

Fig. 2: Performance of German subjects in condition B1 unspliced
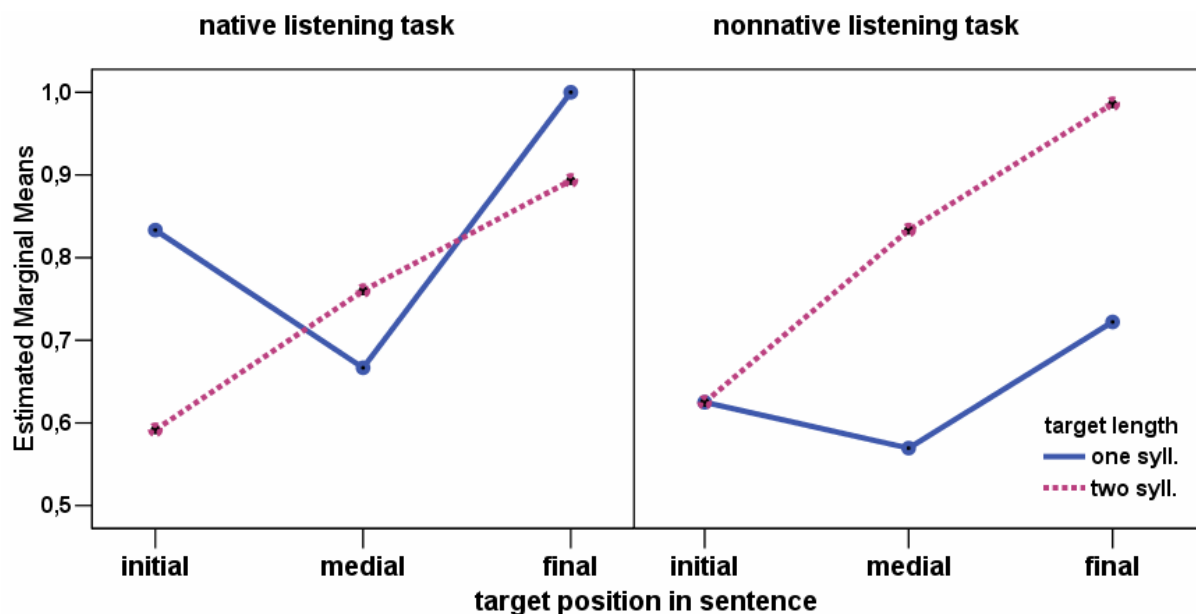
In the condition with an accented target (N1), target length did not have an effect on word recognition in either task. In the native listening task, target

position was significant [$F(2,38) = 5.209$, $p = .01$] and interacted with word length [$F(2,38) = 4.358$, $p < .05$]. Paired comparisons revealed a significant difference between initial and final target position [$F(1,19) = 13.470$, $p < .005$]. There was no effect of target position and no interaction with word length in the L2 task. It appeared that the two patterns of responses in condition N1 are differed regarding the final position in the L1 and L2 task (Fig. 3). Final position does not seem to draw additional attention to targets in second language processing whereas this cues better recognition in the first language.

Fig. 3: Performance of German subjects in condition N1 unspliced
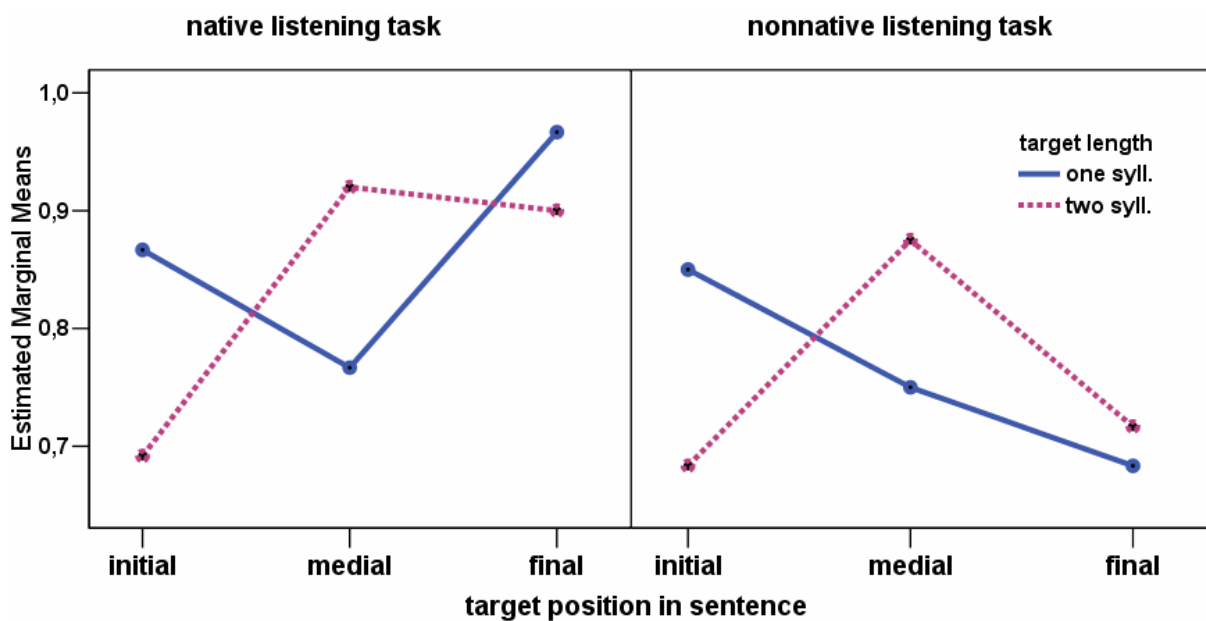
In condition B2 (accent on non-target word in constituent that has the target) there was no effect of word length in the two tasks. In the native listening task, target position was significant [$F(2,38) = 3.261$, $p < .05$] and interacted with word length [$F(2,38) = 4.166$, $p < .05$]. Paired comparisons revealed significant differences between initial and final position of the target [$F(1,19) = 6.265$, $p < .05$] and also between medial and final position [$F(1,19) = 4.388$, $p = .05$]. In

the L2 task the effect of target position was significant $[F_{(2,38)} = 5.028, p < .05]$ and paired comparisons showed a significant difference between the initial and final position $[F_{(1,19)} = 14.241, p < .005]$. There was no interaction of target position with word length in the L2 task.

The response patterns for condition B2 unspliced are shown in Fig. 4.

Fig. 4: Performance of German subjects in condition B2 unspliced
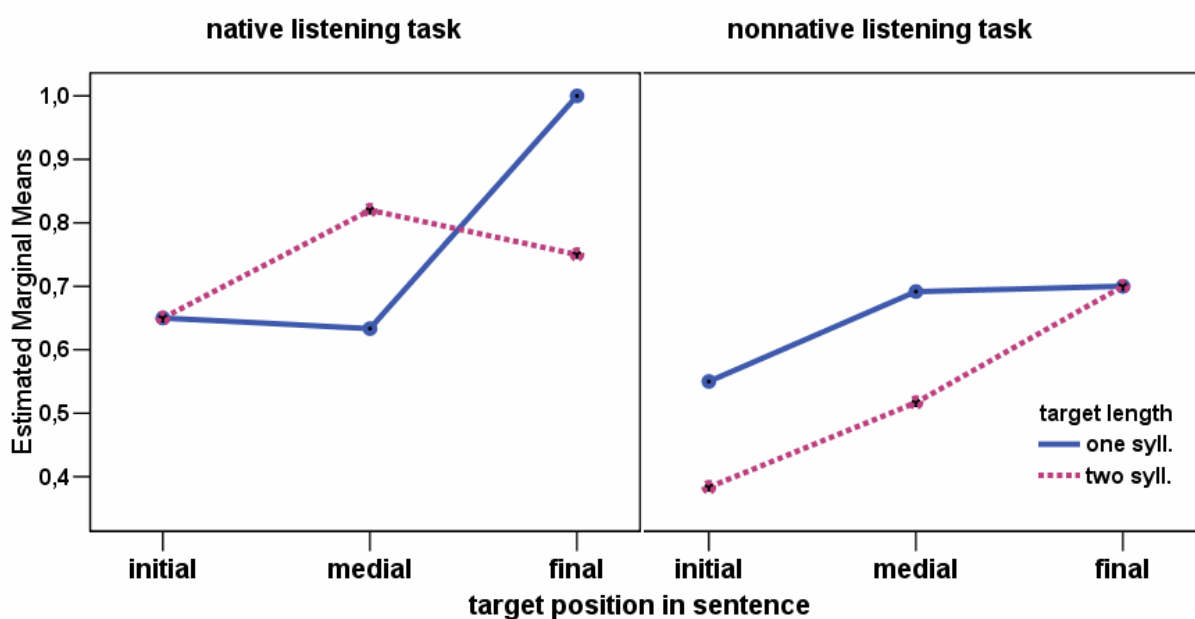
Accuracy scores of the spliced conditions were subjected to the same analyses as described above. In the condition B1 spliced (target word of a B1 broad focus sentence spliced into the context of a N1 narrow focus sentence), there was a significant effect in the L1 task for both target length $[F_{(1,23)} = 26.763, p < .001]$ and for target position $[F_{(2,46)} = 3.253, p < .005]$. There was no interaction between the two factors. In the L2 task of condition B1 spliced, only target position had a significant effect $[F_{(2,46)} = 5.153, p = .01]$ and for this, paired comparisons revealed significant differences between initial and final position $[F_{(1,23)} = 8.013, p < .01]$ and between medial and final position of the

target $[F(1,23) = 10.903, p < .005]$. The patterns of responses in condition B1 spliced are similar in the L1 and L2 task, as can be seen in Figure 5.

Fig. 5: Performance of German subjects in condition B1 spliced
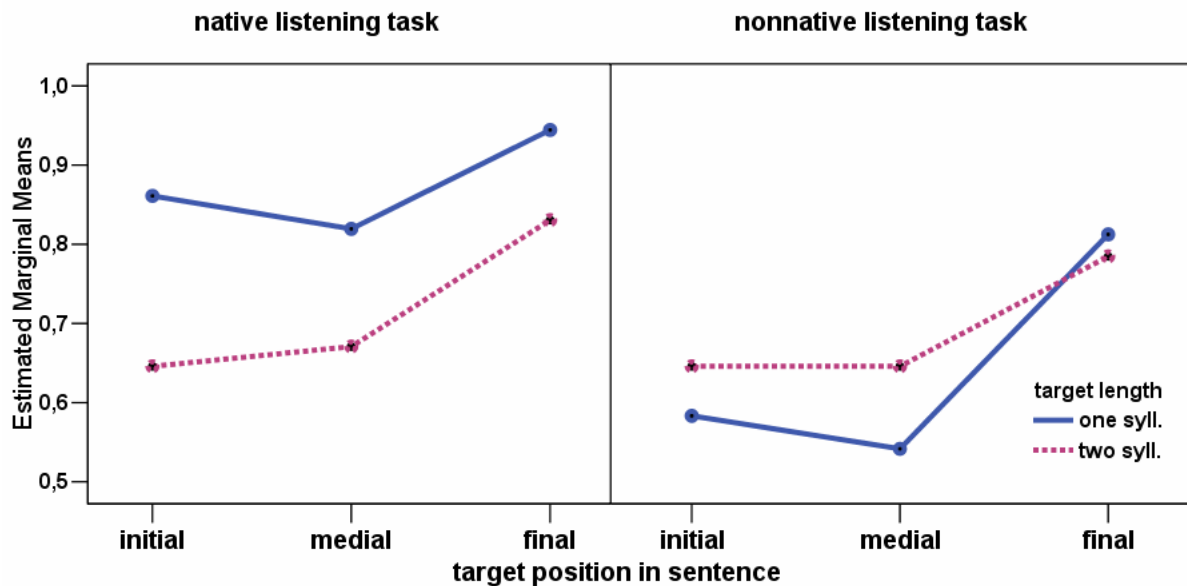
In condition B2 spliced, the target had been spliced out of a B2 sentence (with accent on another word in the constituent that contained the target) into a N1 sentence. In the L1 task there was a significant effect for word length $[F(1,19) = 5.391, p < .05]$ and word length interacted with target position $[F(2,38) = 6.027, p = .005]$. In the L2 task, only target position had a significant effect $[F(2,38) = 3.234, p = .05]$. This was due to a significant difference between the medial and final position $[F(1,19) = 6.883, p< .05]$ (see Fig 6).

Fig. 6: Performance of German subjects in condition B2 spliced
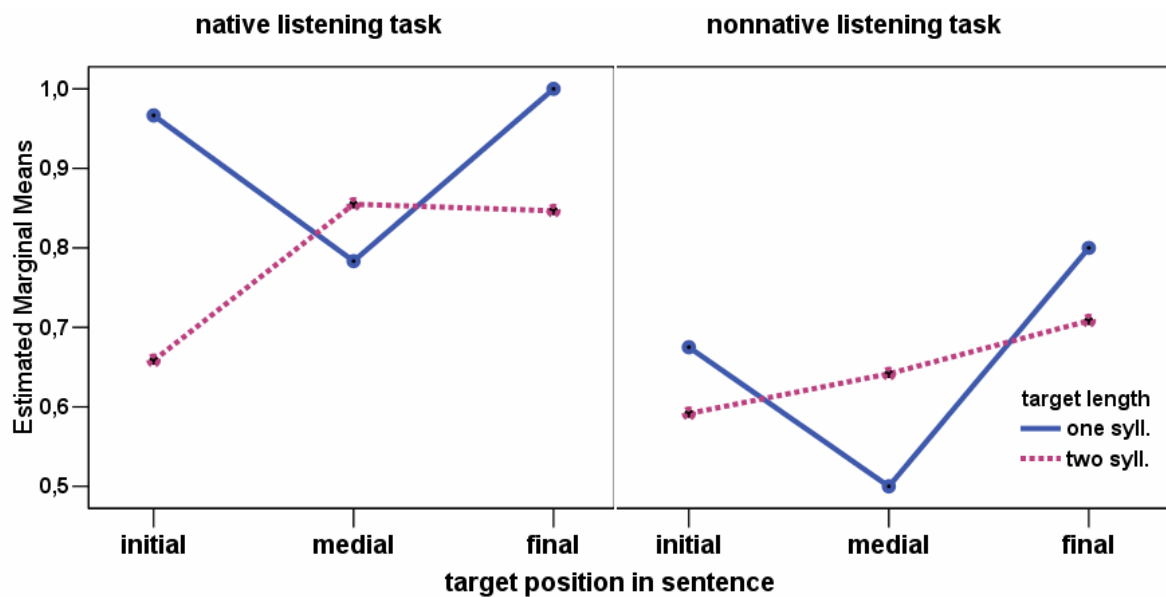
It could be argued from the results that the effect of position found for initial and final position obliterates any possible effect of focus. We therefore examined the recognition accuracy of targets in medial position, thus excluding a superimposed effect of sentence position. Irrespective of word length a stronger effect of focus condition was revealed for targets not occurring at the outer ends of the sentences, where a position effect seems indeed to overshadow focus effect.

Tab. 2: Scores (% correct) per focus condition and part of experiment for targets in medial position

German subjects	B1	N1	B2
L1 task	88,5%	92,2%	88,3%
L2 task	78,9%	90,5%	77,6%
Controls	no values	92,1%	84,2%

3   Discussion

It turns out that accent did not help our subjects process the accented word probes more accurately or more efficiently in L2. There was no significant effect for focus condition. This result confirms the findings of Akker & Cutler (2003) and of Pennington & Ellis (2000). Processing a second language might take up so many resources that accent alone is not a strong enough help. The fact, however, that accent did not help listeners in their L1 either is surprising. This might be due to an effect of target position which is possibly stronger than an effect of accent. This is suggested by the fact that target position had an effect in all accent conditions.

The most salient results were the effects of the position of the target. Over all conditions it seems that position is the strongest cue for the word probe detection task: words in initial or final position of the sentence were remembered better and this confirmed the primacy & recency effect found in other studies (Klein). Our results indicate that in order to effectively commit a representation to memory, the position of the word in the sentence is more important than whether or not a word is accented. It also suggests that effective and efficient processing of accented words (Cutler, 1976) in L1 does not necessarily result in successfully committing the word to memory.

We had hypothesized that learners will be most sensitive to words which are prosodically marked for focus regardless of its location in the sentence and had expected, that German L2 learners of English will recognize target words better when these occur at sentence initial or final position. The effect of sentence position of the target was nevertheless clearly more strong than that of focus accent, which seemed to be merely supportive. Examining the results for target words in sentence medial position only, focus appeared to have indeed a

stronger effect.

The splicing procedure revealed no effects for the cues in the prosodic contour surrounding an accented word and the two test conditions did not lead to conclusive results. In the spliced conditions, target position had a significant effect in L2 and L1.

4 Conclusions

In this study, an experiment was conducted to test which aspects of prosodic prominence facilitate word learning in native and nonnative perceptual language processing. As predicted, words under the narrow focus condition tended to be better recognized than words in broad focus condition.

The results obtained in the spliced conditions confirm findings of earlier studies (Akker & Cutler, 2003), as no effects could be established for the cues in the prosodic contour surrounding an accented word.

Results indicate partly different processing strategies for the two language settings L1 German – L2 English.

The probe detection task used in this experiment may have been a more accurate test of the influence of accent, target length and target position on committing a representation to memory than on accuracy and effectiveness of processing. We think that a phoneme detection task is well-suited to test the effectiveness of processing, and are, indeed, carrying out such an experiment.

Reference

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. Bilingualism: Language and Cognition, 6(2), 81-96.

Allan, D. (2001). Oxford placement test 2: Oxford University Press.

Barcroft, J., & VanPatten, B. (1997). Acoustic salience of grammatical forms: The effect of location, stress and boundedness on spanish l2 input

processing. In A. Perez-Leroux & W. R. Glass (Eds.), Contemporary perspectives on the acquisition of spanish: Production, processing, and comprehension (pp. 109-121). Somerville, Ma: Cascadilla Press.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. Perception & Psychophysics, 20, 55-60.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. Language and Speech, 40, 141-201.

Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. Cognition, 7, 49-59.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. Journal of Experimental Psychology: Human Perception & Performance, 14, 113-121.

Klein, W. (1992). Zweitspracherwerb. In (pp. 78ff.). Frankfurt/M.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for english sentences with prosodic cues. The Modern Language Journal, 84, 372-389.

Rooth, M. (1992). A theory of focus interpretation. Natural Language Semantics, 1.

Rooth, M. (1996). Focus. In S. Lappin (Ed.), The handbook of contemporary semantic theory. Oxford: Blackwell.

Anke Sennema, Ruben van de Vijver, Susanne E. Carroll, Anne Zimmer-Stahl
Universität Potsdam
SFB 632 „Informationsstruktur"
Postfach 601553
14415 Potsdam
Germany
sennema@rz.uni-potsdam.de

Structuring Information through Gesture and Intonation[*]

Stefanie Jannedy & Norma Mendoza-Denton

Humboldt Universität zu Berlin & University of Arizona

Face-to-face communication is multimodal. In unscripted spoken discourse we can observe the interaction of several "semiotic layers", modalities of information such as syntax, discourse structure, gesture, and intonation. We explore the role of gesture and intonation in structuring and aligning information in spoken discourse through a study of the co-occurrence of pitch accents and gestural apices. Metaphorical spatialization through gesture also plays a role in conveying the contextual relationships between the speaker, the government and other external forces in a naturally-occurring political speech setting.

Keywords: Gesture, Intonation, Spoken Discourse, Narrative Structure, Political Speech, Affect.

1   Introduction

It has been widely accepted that the gesture and intonation systems correlate, both aiding in the structuring of verbally rendered discourse (Cassell, 2000;

Loehr, 2004). Yet most of the studies on co-speech gesturing analyzed in the literature result from experimental elicitation and give rise to gestures that are narrative and/or descriptive in nature. That is, participants in laboratory settings are asked to describe something rather concrete such as fish-trapping constructions (Enfield, 2003), or to narrate preselected cartoon strips or films (McNeill, 1992). In this study, we analyze a video recorded sample from a corpus of spontaneous, naturally-occurring data gathered at public Congressional Town Hall Meetings (THMs) in Tucson, Arizona, in which a speaker engages in political discourse, a task much more abstract and goal-directed than elicited narrative, and one in which the speaker is trying to put a political viewpoint across. The speaker (whom we call by the pseudonym *Mary-Jane*) is a woman from Arizona who appears to be in her early forties; her primary interlocutor (U.S. House of Representatives Congressman Jim Kolbe, Republican, Arizona 5th district) is a middle-aged man in his late fifties. They stand in a constituent-representative relationship to each other in the United States political system. In this THM, the co-speech gesturing deployed by Mary-Jane is used not only to describe the political landscape as she sees it, but to persuade, cajole, shame, provide evidence and otherwise convince the interlocutor and the audience to adopt her point of view. Our findings indicate that gesturing not only correlates with grammar and supports the structuring of the information that is acoustically rendered; at the same time, use of the visuo-spatial field conveys information on the relationship posited by the speaker between the government, its constituents and outside forces.

Kendon (1996) thinks of gestures as a 'spill-over' effect from the effort of speaking. For him, gesture is a separate and distinct mode of expression with its own properties which can be brought into a cooperative relationship with spoken utterances and used in a complementary way. He suggests that a study on how phrases of gesture and phrases of speech are related would throw useful light on

how information is structured. Bolinger (1986, p. 199) proposes that "gesture and speech/intonation are a single form in two guises, one visible and the other audible", and argues that gesture and speech stem from the same semantic intent. McNeill et al. (2001, p. 23) argue that "the organization of discourse is inseparable from gesture and prosody" and that "[they are] different sides of a single mental communicative process" (cf. Enfield, 2003, p. 45-46). The results of McNeill's experimental studies indicate that motion, prosody and discourse structure are integrated at each moment of speaking.

Languages such as English or German, unlike for example Swahili (McNeill, 1992), use few morphosyntactic cues to structure discourse. Instead, information is structured primarily by syntactic and prosodic means. Vallduví and Engdahl (1996) discuss how information is packaged into prime units, and give evidence for the difference between English (which uses primarily intonation) and Catalan (using primarily syntax) in the packaging of information.

The role of gesturing during speaking and its role in the structuring of discourse has until recently been largely unexplored. Yet the relationship of gesture to speech increasingly captures the interest of human-computer interface designers aiming to model the movements of animated agents, since it has been shown that gestures facilitate information comprehension (Beattie & Shovelton 1999). Gesture researchers have further shown that subjects can recall information that was selectively presented in the gestural but not in the verbal channel (Kelly et al., 1999; Cassell et al., 1999; McNeill, 1992), all without being able to remember what channel the information was presented in. These results are provocative because they imply that as far as the design of animated agents is concerned, greater information density can be achieved by presenting part of the information acoustically and part of it (possibly complementary or reinforcing information) visually, bringing us closer to modelling the online

workings of face-to-face conversations. A thorough description of gestural-speech co-dynamics is thus necessary for an algorithmic approach to be formulated, and essential in the design of automated agents and in achieving some naturalness in their movements (Cassell, 2000 and references therein).

The purpose of this research is to describe and understand the timelines involved in a case study of the complementary presentation of information. Part of the information is presented through the speaker's verbal stream, while the broader and complementary political setting and the assumptions on which it rests are presented gesturally. We have independently verified that both of these meanings come across by asking audiences of graduate and undergraduate students in English-speaking universities (at the University of Arizona, the Ohio State University, and at University College Dublin in Ireland) to describe what they think the speaker was trying to get across. Three different modes of presentation of our data have been judged by the student populations:

1) Acoustic and visual information: Audiences who have seen the video and listened to the audio have repeatedly described the complementary information presented on both channels.

2) Acoustic information only: Audiences who have been presented with the audio portion describe the main assertions presented in the spoken discourse with no reference to the larger political landscape.

3) Visual information only: Audiences who have only seen the video without the audio merely impute anger and emotionality to the speaker.

These results highlight one well-known fact about the relationship of gesture when it is ancillary to spoken discourse: while spoken discourse has a high

referential resolution, that is, it is able to pick out referents with relatively little ambiguity, gesture has a low referential resolution, so most of the information presented gesturally is supportive to speech and not recoverable solely from the gestural channel (an obvious exception to this being gestural systems that are full-fledged languages, such as American Sign Language). Gestures and spoken discourse are thus complementary in nature, and have different affordances and limits to their ability to present information. Harper et al. (2000) have found that in a study of 3D multiplayer wargame interactions, gesture was used for much more than just simple deictic functions. They explain:

> …language facilitates complex queries with the ability to express quantification, attribute and object relations, negation, counterfactuals, categorization, ordering, and aggregate operations. Gesture is more natural for manipulating spatial properties of objects (size, shape, and placement) in graphical environments. (Harper et al. 2000:3)

In Harper et al.'s study, subjects participated in a wargame simulation which involved some players working around a scaled, 3D model of a battlefield, while other players were consulted remotely. In post-game discussions, participants remarked on the difficulty of communicating with those not working around the 3D model being used (2000:5). Because of their heavy reliance on definite descriptions in combination with gestures in the face-to-face game, players encountered difficulties when they were restricted to using only the verbal channel.

We hold there to be a parallelism between gesture and speech, both of them carrying meaning on at least three different planes: structure, content, and social meaning.

The acoustic signal is an enormously rich source of structured information. from a phonetic-phonological point of view, there are positional restrictions or co-occurrence restrictions of phonemes, the phonotactics that make up a specific language variety. There is a specific prosodic or rhythmic structure, such that some parts of a word are uttered louder and longer lending a specific syllable more perceptual prominence either to indicate lexical stress or informational salience. At the same time, obviously, the acoustic signal also contains other grammatical information. That is, lexical or semantic content is being transported by the choice or words spoken, and pragmatic content is chosen by the context the words are uttered in. Further, the acoustic signal transmits social content, that is, it also carries social information pertaining to the speaker (age, gender, ethnicity, etc.) and their interlocutors.

Co-speech gestures correlate with grammar, that is, they correlate with grammatical structure by indicating beats (correlated with pitch accents), or gestural phrases (correlated with syntactic phrases). These same gestures also carry semantic and pragmatic content by being deictic or emblematic or just emphatic beats on parts of the information emphasized by the speaker. Simultaneously, these same gestures also transmit social content in terms of how the speaker recounts and relates to the world or abstract universe surrounding them.

The gestural information stream is able to work online and incorporate context to take advantage of props and of the spatial location of the audience. We argue that the communicative constraints of the sociolinguistic situation in our case study maximized the need for the simultaneous presentation of information, since the speaker was only grudgingly given the floor in the first

place, was constrained in the time she was allotted for her turn at talk; while the interlocutor constantly threatened to interrupt her, attempted to shift his attention away from her and to cut her off. Thus, the sociolinguistic pressure was great for the speaker to pack as much information as possible into her short turn at talk. For our purposes, this results in a naturally-occurring situation where time and interactional constraints push the limits of both intonational and gestural information packaging.

2   Background

2.1   What do hand, arm, and mouth movements have in common?

Both the motor theory of speech production/perception (Liberman and Mattingly 1985) and the articulatory phonology account of speech as a stream of articulatory gestures (Browman and Goldstein, 1990, 1992) place articulatory movements produced during speech and "paralinguistic" gestures on the same cognitive plane, as they are both thought to be *coordinated patterns of goal-directed articulatory movements* (Tatham 1996). These approaches require that speech movements and gestural movements be accounted for by the same mechanisms, since it is assumed (contra Chomsky 2000, Fodor 1983, Marantz to appear) that the structures of language are not modular or unique in comprising a "language organ", but rather that these structures were derived and modelled on the pre-existing neural systems which had evolved for the control of body movement. Rizzolatti et al. (1988, 1999) posited a class of premotor neurons, the "mirror" neurons based on their experiments with monkeys. They explain: "…with this term [mirror neurons] we define neurons that discharge both when the monkey makes a particular action and when it observes another individual (monkey or human) making a similar action. […] Transcranial magnetic stimulation and positron emission tomography (PET) experiments suggest that a

mirror system for gesture recognition also exists in humans and includes Broca's area (1988:92)".

The results of these earlier studies of motor action have been replicated, and links shown between the sensorimotor system and the acoustic system both in monkeys and in humans (Berthoz, 1997: Fadiga & Craighero, 2003). The relationship between speech and limb movements is thus taken for granted, since both derive from elementary motor programs or *articulatory gestures*. According to Berthoz (1997:176) "Le mouvement est donc organisé à partir d'un répertoire de synergies qui compose autant d'actes possibles…une bibliothèque de mouvements facilement déclenchables. [*Movement is then organized beginning with a synergistic repertoire which makes up possible actions…making up a library of movements which are easily carried out.]*"

Both speech gestures (Browman & Goldstein, 1990) and hand/arm gestures then are finely coordinated and stem from higher level cognitive processes through which information is structured. While there is no one-to-one correspondence between form and meaning in gesturing, gestural organization maintains a tight link with the semantics of speech. That is, gestures strongly correlate with grammar and grammatical structures: it was found that the stroke (peak of effort) of a gesture occurs with the intonationally most prominent syllable of the aligned speech segment (Kendon, 1980; McNeill, 1992; Cassell, 2000; Loehr 2004).

The significance of the motor theory of action for our study is the following: we believe that when a speaker executes a particular gesture, the embodied nature of the gesture causes the interlocutor to generate an internal representation of the movements in order to decode and interpret the spatial field (concrete or abstract, see our discussion of Krifka 2005 below) depicted by the speaker. This is essentially a parallel to what the motor theory holds for speech (Liberman and Blumstein, 1990:147), where we practice "analysis by synthesis"

(Halle and Stevens 1959) in generating an internalized representation of the incoming signal.

Thus, given an interlocutor that shares spatial conventions, speakers may gesture and have their concrete and abstract meaning understood without uniqueness of referents. Spatial conventions exhibit cross-cultural variability: Haviland (1993), for example, has found that narrative retellings involving co-speech gesturing in the Gugu Yimidhirrh aboriginal language spoken in Queensland, Australia, exhibit absolute directionality, that is to say, when a speaker retells an incident their deictic pointing will refer to the same cardinal direction in which the relevant incident took place. Similar findings have been reported for Amerindian Languages such as Tzotzil (Brown and Levinson 1993) and Assiniboine (Farnell 1995). Other speakers, especially those from Western cultures, exhibit relative directionality, so that their gestures tend to be located relative to the ego (though some flexibility also exists here, with speakers able to present gestural situations from different points of view (McNeill 1992:190). There may additionally be abstract topic/comment structuring in the directions of pointing or in the handedness of a gesture (Krifka 2005). We hold that topic/comment structures will be present in preferential hand-gesturing, but will be forced to interact with culturally-conventionalized gesturing directionality. This would predict that, given the same story stimuli, speakers from absolute-directionality gesture traditions would encode topic/comment gestures differently from those of ego-centric gesture traditions. We exhort the next generation of linguists and linguistic anthropologists to conduct these experiments.

2.2   A typology of gestures

To gain a better understanding of the classification of gestures, a short digression into gesture theory seems necessary. From a kinetic point of view, gestural movements are described to have obligatorily three and at most five phases (McNeill, 1992): The *preparation phase* (optional) marks the beginning of the motion in which the parts of the body involved in the gesture leave the neutral position and move to the position necessary for the upcoming gesture. The *pre-stroke hold* (optional) is the position of the hand/arm at the end of the preparation phase before the beginning of the stroke. The *stroke* (obligatory) is the climax or peak of effort of the gesture. It is one of the most recognizable components of a gesture, it is synchronized with the linguistic forms, such as accented syllables it is co-expressive with. The *post-stroke hold* is the position the hand/arm remains in when the co-expressive spoken utterance is delayed. The *retraction* is the end of the motion in which the parts of the body involved in the gesture return to neutral positions. The *beat* gesture is smaller and is described to only have 2 phases which are typically flicks with the wrists or fingers. In the case of Mary Jane, we will see that her impassioned argument motivates her marking beats with her entire upper torso (Panel 4.2).

From a semantic or pragmatic point of view, gestures can have different attribute values:

1) *Deictic* gestures point towards a concrete or abstract referent;
2) *Beats*, sometimes called *batonic gestures* because they resemble a conductor keeping an orchestra in time, are rhythmic gestures that have no specific form but which are synchronized with the speaker's utterances.
3) *Iconic* gestures depict a concrete object or event in a narrative;
4) *Emblematic* gestures replace speech (for instance, shaking one's head to mean "no").

5) *Metaphoric* gestures represent an abstract idea.

McNeill (1992) calls the recurring combination of the same gestures with prosody and discourse organization *catchments*. These catchments are recognized from recurrences of gesture form features over a stretch of discourse (two or more gestures with partially or fully recurring features of shape, movement, space, orientation, dynamics etc.) and serve to offer clues to a cohesive linkage in the text in which it occurs.

Rather than assuming that gesturing only serves the interlocutor in structuring acoustically rendered information, we take it that gesturing aids both the speaker and the interlocutor: recall Harper et al.'s (2000) wargame simulations referenced above. It has also been observed that gestures occur no less frequently when talking over the telephone and the speaker cannot be seen by the listener (Cosnier 1982; Rimé 1982).

We also assume that the very same gestures that are so tightly aligned with grammar and prosody serve as a device to mediate between the speaker and the world. What does this mean? By having gestures that are ego-centered, and placing herself in the middle of a depiction of the political landscape, Mary-Jane sketches *through gestures* her view of an idealized public sphere, and of the relationship between a constituent and the broader powers of government.

2.3   Intonational Grammar

The grammar of intonation we are assuming goes back to Pierrehumbert (1980) and was formalized as a transcription system for prosodic annotation summarized in Beckman & Ayers (1994). According to this grammar of intonation there are pitch accents, intermediate phrases and intonation phrases. The grammar of intonation can be summarized as follows: each intonation

phrase (marked with '%' at the right edge) must contain at least one intermediate phrase (marked with a '-' at the right edge), which in turn must contain at least one pitch accent (marked with a '*').

Pitch accents are associated with the stressed syllables of words. These accents mark local prominences above the level of the word in an utterance. There is a fixed inventory of pitch accents, they can have different tonal shapes that are marked with labels such as H*, L* or L+H* etc. The '*' indicates that there is a pitch accent which is defined as a tonal target. The tonal events between pitch targets are accounted for by interpolation between these pitch targets. For example, one should observe a falling fundamental frequency contour when a H* tone target is followed by a L* tone target. Downstep is a phonologically triggered process by which a H* accent is downstepped (!H*) when it occurs subsequent to the downstep trigger. Often this trigger is a bitonal pitch accent type such as L+H* or L*+H accent. By definition though, the domain of downstep is the intermeditate phrase, thus, the trigger must occur in the same intermediate phrase as the downstepped accent. Downstepping is a compression of the pitch range lowering the F0 targets for the H* accents following the downstep trigger.

An intermediate phrase is a minor phrase and consists of one or more pitch accents plus a phrase accent associated with the right edge. This phrase accent can be either L or H. An intonational phrase consists of at least one or more intermediate phrases. The right edge of the intonational phrase has a phrase tone, taking the shape of L or H. These edge tones determine the shape of the F0 contour between the last pitch accent and the end of the phrase.

3   Data and Methods

We utilize video data collected in 2000-2001 for an ethnographic study of Congressional Town Hall Meetings (THMs) in Tucson, Arizona. The data forms part of a fieldwork project on political discourse, language and power. A total of approximately 20 hours of data was collected at locations around Southern Arizona and Washington, D.C. THMs are public fora announced in the media, on websites and through fliers to homes in the neighborhoods that are represented by Kolbe. A THM in this district normally lasts one and a half hours to two hours and is led by Congressman Kolbe, typically taking the form of an initial period of question-writing by the audience on slips of paper circulated by Kolbe staffers. This is followed by a rehearsed monologue from Rep. Kolbe in which he states as his aim the updating of his constituents on important happenings in Washington, D.C. Then Kolbe selects some of the questions that constituents have written out as questions to be answered without calling on anybody specifically. After this he might take a couple of spontaneous questions, run to the end of the allotted time and then invite those who wanted to talk to him further to stay and discuss matters after the official THM ends.

The video data of Rep. Kolbe and Tucson citizens that we analyze for this paper was recorded simultaneously with 2 video cameras, both located to the left of the audience (from their perspective), one pointed toward Rep. Kolbe and the other more aimed more generally toward the audience. Rep. Kolbe was miked with a lavaliere microphone plugged into camera 1, and the audience sounds was captured by a microphone mounted on camera 2. By aligning the sound tracks of the two videos, we were able to synchronize the videos exactly so as to gain accurate descriptions of the hand and arm movements from two different angles. Both camera 1 and camera 2 capture the audience from different vantage points.

On this particular occasion, on a Saturday morning in February 2001, the THM was held in the cafeteria of a midtown school. The congressman introduced the researchers and advised participants in the THM that they were being video taped for a research project and that their participation was strictly voluntary. If any of the constituents objected to being taped, they were encouraged to approach the researchers after the THM. None did.

The data we selected for microanalysis for this study lasts exactly 130 seconds (a reasonable amount of data within the gestural analysis literature (See Loehr 2004 for a discussion)), and was transcribed according to the ToBI intonational transcription framework in addition to being subjected to a modified McNeill-style gestural transcription. Eleven tiers of body movement and gestural transcription were coded.

3.1   ToBI and Gesture Transcriptions

Trained linguistics and anthropology students at the University of Arizona worked on transcribing the data on several different levels by using the program Praat. Orthographic transcriptions were done and then checked by another student. Prosodic transcriptions were made in a team of three students of linguistics by majority vote: They were trained to transcribe intonational events within the ToBI (tone and break index) framework based on the Pierrehumbert (1980) system of English intonation. In cases of uncertainty or disagreement, they discussed the issue until a consensus was reached. Then, all prosodic transcriptions were checked by an experienced labeler of US-English (one of the authors). It must be noted though that the data was very difficult to annotate due to the fact that the recordings were made in a naturalistic setting. In addition to room noises (random background noise) which shows up in spectrograms as energy in all frequency regions, the audio data we analyzed contains claps and

coughs from members of the audience, partial interruptions by the congressman and other noises that are unidentifiable. In figure 1 below, taken from our Praat display, we show the sound pressure wave (waveform) in the upper display, and a spectrogram (individual frequency bands) which was overlain with the fundamental frequency trace (F0) necessary for the tonal analysis. This is how we have coded lines 23-29 of the full transcription included as an appendix to this paper.
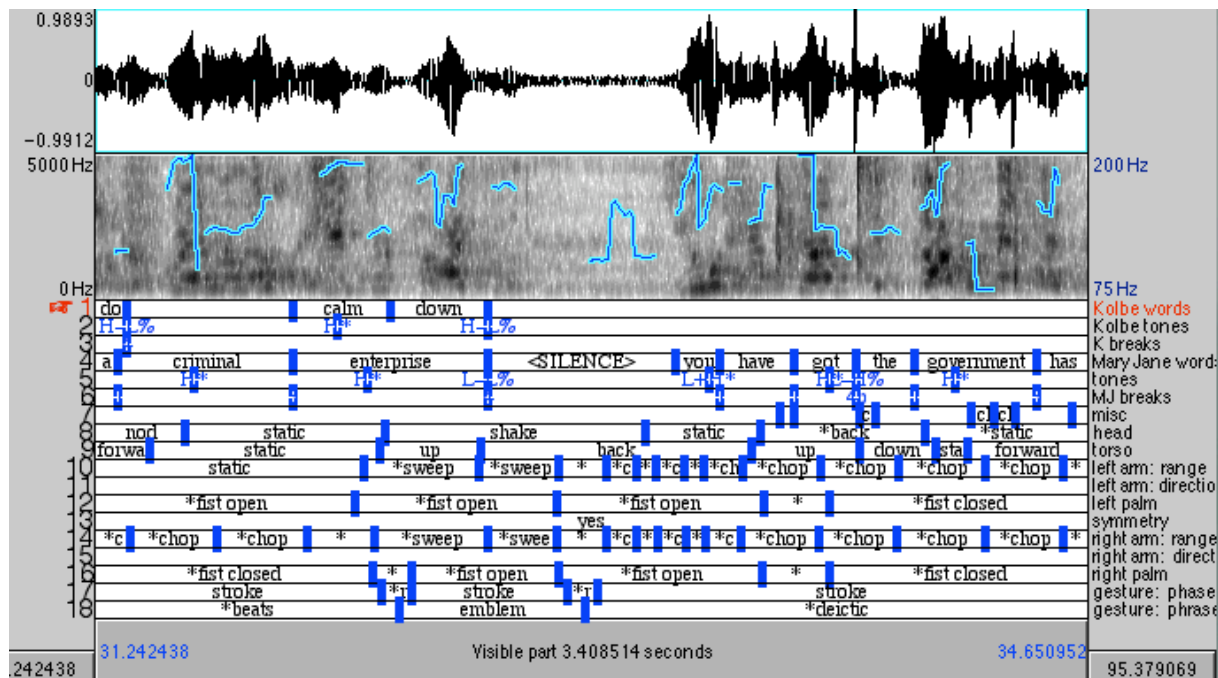
Figure 1: Praat display of multitier transcription of intonation and gesture.

Though there is a growing body of literature on gesture, no standard transcription system for manual gestures has been agreed upon and described. There are a couple of systems which are in relatively wide use; within language studies, the most prominent is McNeill's system, developed expressly for gestures. Based on gestural primes proposed by McNeill (1992), we transcribed movement of the left and right arm and the left and right hands on 6 tiers (range

of movement, direction of movement, palm configuration for each side). In addition, head and torso movements were transcribed as well as whether or not there was a symmetry in the movements of both hands and arms. We also annotated the gestural phases which have kinetic properties (preparation vs. stroke, for example) as well as the gestural phrases which are more semantic or pragmatic in nature (deictic vs. metaphoric etc.).

In order to have an exact alignment of audio and video, the data was looked at in ANVIL, a program that allows video and audio to be time aligned for transcription purposes. For the purpose of our study it became crucial not only to mark intervals during which gestural movements were performed, but to mark points as well. We were faced with the issue of data reduction since it seemed impossible to gain any insights from the amount of data available to us. As an initial step, we wanted to investigate if and when and how often pitch accents would co-occur with gestural movements. Since we had mainly coded intervals, and pitch accents are by default point events in time, we added another transcription level in which we coded what Loehr (2004) has called the *apex*.

The apex is the point in the hand or arm gesture during which (in Task Dynamics terms) the equilibrium of a particular gestural movement is reached (Browman & Goldstein, 1990). Task Dynamics holds that gestures are defined in terms of tasks (extend arm, point with finger etc.) which are the coordinated movements of several limbs. Further, the gestures are defined in terms of the dynamics that specify the motions in terms of a mass-spring model applied to articulation. Gestures are defined in terms of three specifications: 1. the stiffness of the gesture, the target of the movement and the phasing of the gestures with regard to each other. The stiffness roughly corresponds to the speed of the gesture, that is, speed is a reflection of the stiffness of the gesture. The stiffer the gesture, the faster it is being executed. When the target is being approached relatively slowly then there is less stiffness. The amplitude of the gesture is the

amount of displacement from the resting position. The phasing of gestures then is the intergestural timing, that is, the timing between gestures, specifying when one gesture begins and when the next gesture ends.

In terms of our specification of the apex, we adapt Loehr's (2004:89) definition. He describes the apex as the "peak of the peak" or as the "kinetic goal of the stroke". In task dynamics terms then this is the target of the gesture. The reason why this is described here in some detail is that we noticed that in more rapid, highly emotional speech, the co-speech gestures are not fully carried out, there often is just a succession of apices, tightly coordinated with pitch accents without the hands or arms returning to a resting position. The gestures overlap and cut each other off, only leaving the peaks of the gestures being observed. (In our example we find that the speaker nods her head for emphasis as well. We have however excluded this from our analysis and annotated hand, arm and body movements only).

4   Analysis

The total length of the sequence we selected for analysis is 2 minutes and 10 seconds. We have provided a full transcription of the whole sequence of Mary-Jane's monologue as Appendix A at the end of this paper. For microanalysis we have selected several segments which most clearly show the various gestural and intonational effects that we aim to illustrate. These segments are subdivisions created by us in our data and are not meant to be compared against each other, but to illustrate sequences that we felt had a unity of topic, sequencing and purpose within the total data set. They are presented in the order in which they occurred, as clusters of frame grabs in panels 1-4. Each panel has been divided into individual frame grabs in order to show the exact gestural sequences. Thus, P2:12-14 means Panel2, frames 12 through 14. We will follow

the following format in the presentation of this data: First we will present the transcript along with associated annotations, then the corresponding panels and finally our discussion and any illustrative figures related to them. Let's begin with the first sequence:

4.1   Panel 1 analysis: The drug war will be over.

M-J	the drug war will be over. **P1:1-3**
(0:20)	we won't have to ship guns and military to Colombia, **P1:4-11**
	it will be over. **P1:12-15**
	Bush won't have to talk to Fox about the drug war, **P1:16-24**
	it will be **P1:25-27**
	over. **P1:28-30**
Audience	[XXclappingXX]
	[there will not be a drug war] **P1:31-36**
	if you legalize drugs. **P1:37-42**

Panel 1: "It will be over"

In P1:5-11, Mary-Jane begins with a deictic gesture with the origin close to her chest. As she utters the phrase, "we won't have to ship guns and military to Colombia," she extends her right arm out and points out to the right, metaphorically placing Colombia to one side and away from the origin at ego.

This gesture has a clear parallel with later gestures, most notably the one in P1:16-24, where M-J does a bimanual pointing gesture immediately in front of her and then extends both arms to the left while saying, "Bush won't have to talk to Fox about the drug war", producing apices and pitch accents aligned with "Bush" and "Fox." Interestingly, though both Colombia and current Mexican president Vicente Fox (and by extension, Mexico) are located south of her, she has placed them to the right and left respectively, while Bush, who is north-east of her, has been placed directly in front of her (she herself is facing absolute south). Although we would not expect a speaker of English to exhibit absolute cardinal directionality, people in Tucson often do, since the city is two hours north of the border with Mexico and has dramatic mountain ranges which orient residents to absolute directions. This apparent disparity in the directions of well-known places was our first clue that something beyond simple deictic direction was being represented through Mary-Jane's gestures. A bird's eye representation of her gestures would show Bush, guns and military directly facing M-J, while Fox and Mexico would be to the left side and Colombia to the right (Figure 2). :
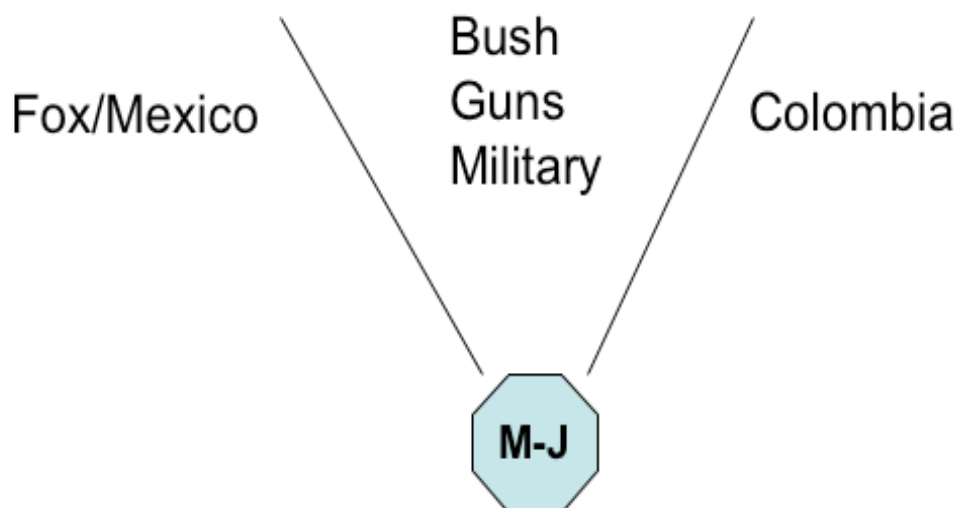
Figure 2: Birds' eye representation of spatialization of entities by M-J, Panel 1

An interesting catchment of gradually exaggerated gestures also occurs in this stretch. M-J repeats the phrase "it will be over" three times, each time with successively downstepped intonation and more emphatic gesturing. The first time, she brings her left hand (which is in ASL handshape G/X1 (McNeill 1992:88) in one downward stroke from eye level to shoulder level (this is visible in P1:1-4). For the second iteration of the phrase "it will be over," she draws a three-sided, open-bottom box with both hands, starting at the center top (P1:12-15; P1:14 shows both palms facing each other as she sweeps them down to make the sides of the box).

In the final iteration of the "it will be over" phrase (P1:25-30), she draws the complete box and slows down her speech, exhibiting an intonational phrase boundary and a gestural hold in the exact spot where she prepares to bring her hands in to close the box (P1:29), right before the last word, "over." The nested structural parallelisms appear thus:

1) lexical, intonational and syntactic parallelisms with coindexed pronouns at ACE;

2) intonational, gestural, and syntactic parallelisms at BD,

3) and repetition along with gestural parallelism and expansion at CE.

A   The drug war will be over

B   We won't have to send guns and military to Colombia

C   It will be over

D   Bush won't have to talk to Fox about the drug war

E   It will be over.

Tannen (1989) has argued that repetition in discourse has a cohesion and focussing function, serving to highlight information and structure listener

expectations. We believe the communicative constraints that we mentioned above motivate the use of four different semiotic levels of repetition (lexical, syntactic, phonological (intonation), and paralinguistic (gesture)), all within a space of nine seconds of spontaneous speech. The fine-grained coordination of moment-to-moment speech has been discussed in the conversation analytic literature, and to some extent in the literature that is called "emergent grammar," which is concerned with finding the emergent structure of speech as it happens in naturalistic interaction. Our description of the incorporation into emergent structure of affect-laden gesture and intonational phenomena contributes to this literature.

4.2   Panel 2 Analysis: Excuse me, hello?

	L+H* H* L-H%
M-J	because they're going to ex**pe**riment with **drugs**, (P2:1-3)
	L* L* L-H%
	like **ev**ery one has **done**, (P2:4-9)
	L* L-H% H* L-L%
	since **day** **one**. (P2:10-14)
Kolbe	ok,
	well I--[
	H* L-L%
M-J	] we ex**pe**riment. (P2:15-16)
	H* H* H* L-L%
	alcohol and to**ba**cco **kill** people (P2:17-23)

```
        H*              H*  H*           L-L%
   marijuana doesn't kill   anybody. (P2:24-28)
       L*+H            L-H%
   excuse me (P2:29-32)
   L+H*         !H-L%
   hello: (P2:33-36)
```


Panel 2: "Excuse me, hello?"

This segment is annotated with the ToBI tones marking pitch accents (*) as well as phrase tones (-) and boundary tones (%). The bolded words indicate that there is a simultaneous occurrence of a pitch accent with a gestural apex. It is particularly noteworthy that Mary-Jane deliberately separates the phrase 'day one' into two intonational phrases. Both words are lent prominence by accenting them and by making pointing gestures where her right index finger lands on the upwards-open palm of her left hand (day: P2:10-11 and one P2:12-14). Note also that Mary-Jane produces the last gestural apex in this segment with her entire upper body as the articulator, which she abruptly stops in mid motion at a precarious 45-degree angle as she says to the Congressman "excuse me, hello".

The intonation contour L*+H L-H% occurring on 'excuse me' has been described as carrying a holistic pragmatic meaning ranging from uncertainty to incredulity (Ward and Hirschberg 1985, Hirschberg & Ward 1992). Hirschberg and Ward argue that "when speakers use the contour to express incredulity, they generally express that incredulity about a value already evoked in the discourse" (p. 243). A striking fact about this use of the uncertainty/incredulity contour is that it is not aligned with the utterance that one might expect it to, given the partially-ordered set relationship (poset) framework described in Ward and Hirschberg (1985). Consider the relevant part of the utterance once more:

       H*   H*   H*   L-L%

alcohol and to**bacco** **kill** people (P2:17-23)

         H*   H* H*   L-L%

mariju**a**na doesn't **kill**   **a**nybody. (P2:24-28)

       L+H*   !H-L%

ex**cuse** me (P2:29-32)

In denying that marijuana kills people, Mary-Jane invokes the possible set of substances that do kill people, and discursively selects for the interlocutor (Kolbe) and overhearers (other constituents) deadly substances that are *legal*. The incredulity being expressed here, signalled by the wide pitch range that the contour carries and which upgrades it from uncertainty (Hirschberg and Ward 1992), is in the assertion that it is the noxious substances that are legal while the innocuous substance is illegal. These discourse facts all line up with the licensed uses of this contour according to Ward and Hirschberg (1985), with one exception: the contour doesn't fall on the expected phrase (marijuana…), but instead on the *following* one (excuse me...). We believe that this finding has two interpretations with important implications:

1) It is possible that this is the point where a speaker under strong communicative demands finally reaches the limits of the processing capacity for online information alignment. In this short segment, Mary-Jane simultaneously aligns pitch accents with syllables contained in informationally prominent words, gestural apices with those pitch accents, while making a complex argument and taking parts of her body to the limits of their physical space. Could the canonical alignment of the incredulity contour have been given up in order to meet the extra processing demands of this task?

2) A second interpretation is that the speaker, knowing that she was going to use a conventionalized expression with its own L*+H L-H% contour ("excuse me, hello?" often is uttered this way colloquially in the United States), chose to suppress the first contour so as to avoid two identical contours together, following a kind of prosodic OCP, or presumably

because the meaning is already accessible from a single utterance of the contour.

On "hello", Mary-Jane produces a pitch contour that we described with the tonal sequence L+H* !H-L%. This contour has also been described as a 'calling contour' (Beckman & Ayers, 1994). It is the contour often used when shouting a name during the process of looking for somebody (for example to call somebody for dinner). It is our impression that Mary-Jane uses this contour to call Congressman Kolbe metaphorically to do his job, to reproach him.

4.3 Panel 3 Analysis: Alcohol and Tobacco

|	
M-J	]why aren't you talking about alcohol and cigarettes? (P3.1:1-7)
	what you've done is you've taken that industry, (P3.1:8-12)
	and you are putting it on third world. (P3.1:13-17)
Kolbe	what--
M-J	they are consuming the tobacco now. (P3.1:18-23)
	that's where all the tobacco money is coming from. (P3.1:24-30)
Kolbe	[we're spending]
M-J	[they're still making] a profit on cigarettes (P3.2:31-37)
	cause they're selling them to all the [people] (P3.2:38-40)
	that we're dumping our chemicals on (P3.2:41-48)

Panel 3.1: "Alcohol and Tobacco."

Panel 3.2: "Alcohol and Tobacco."

The segment represented by panels 3.1 and 3.2 returns to the earlier metaphorical spatialization described in Figure 2 despite the intervening segment in panel 2. Also before panel 3, Mary-Jane has been referring to "kids you've been talking about today" who will go to prison because of their alcohol and drug use. Interestingly, the gesture in "kids" starts out directly in front of Mary-Jane, metaphorically in a space shared between the government and the people, and after drug use the "kids" get moved off to "prison," which also occupies the peripheral position earlier used for external entities (other countries) outside the relationship between the government and the people. Another movement from "us" to "them" takes place in Panels 3.1 and 3.2: as Mary-Jane discusses the practice of chemical dumping by the United States, she makes a gesture that starts out with both hands right behind her left shoulder (as though she were holding a ball and beginning to toss it, P3.1:10), and winds up gesturally in what is by now in her gesture system a stable place for the "third world," to the far right in her spatial system P3.1:18.
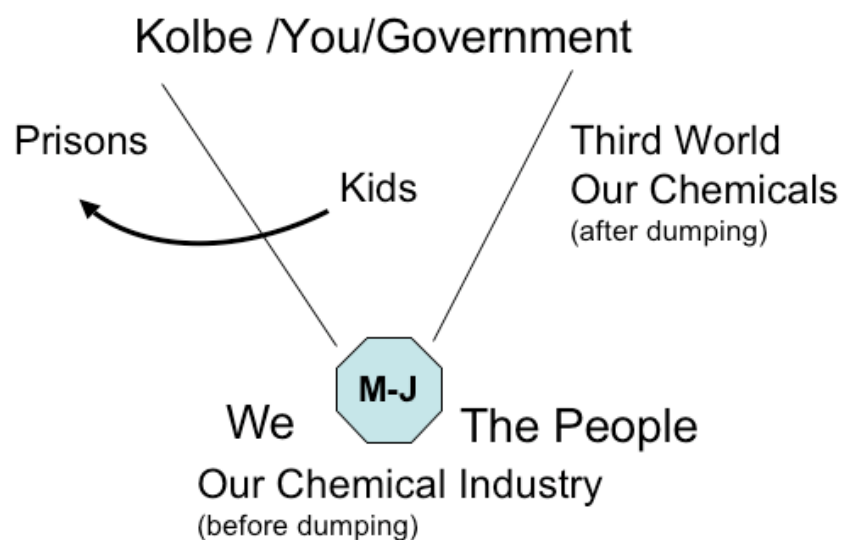
Figure 3: Bird's eye view of spatialized entities represented by M-J, Panels 3.1-3.2

Although a detailed discourse analytic account of the shifting pronominal referents in this data is beyond the aims and scope of this paper, we will note that the pronoun "we" variously refers to a) the people of the United States as a collective that produces chemicals that need to be dumped someplace ("people that we're dumping our chemicals on"); b) the United States as an entity that ships guns and military overseas ("we won't have to ship guns and military"); c) constituents that engage in collective behavior and do not share in the government's definition of what is legal and illegal ("we experiment"); and d) individual citizens who have rights to control their own bodies ("yes we do.") The pronoun "you" likewise shifts in reference to the addressee in general, to representative Kolbe, to current president G.W. Bush, and to the government in general. This variability in the instantiation of pronominal reference has been observed in political discourse cross-linguistically, as political figures (Van Dijk 2003, Wodak 1989) act on the political stage. In this case tracking the shifts in pronominal reference helps us to anchor and interpret the gestural data, and to make explicit the implicit political model that Mary-Jane sketches with her gestures.

4.4   Analysis Panel 4: My rights as a citizen

M-J   they want to make (P4.1:1)

(.)

hemp seed (P4.1:2-4)

a schedule one narcotic (P4.1:5-11)

(.)

when did I lose the right,(P1:12-17)

(.)

to eat, (P1:18-22)

(.)

the food, (P1:23-24)

(.)

I want, (P1:25-27)

(.)

to eat. (P1:28-30)

when did I lose my rights

as a citizen of this country, (P2:31-41)

to put into my body, (P2:42-45)

and listen to whatever, (P2:46-48)

or watch whatever I want. (P2:49-54)

In the sequence "to eat, the food, I want, to eat" is composed of 4 separate intonational phrases. Each of these phrases contains one intermediate phrase which in turn contains one pitch accent. There are accents on "eat", "food", "want" and "eat". We have observed that perceptually, these accents appear to occur in a downstepping relationship to each other (Liberman & Pierrehumbert, 1984). According to the tonal labeling conventions (ToBI) though, this would be infelicitous. Unfortunately, measurements of the fundamental frequency contour during the time point of the F0 maximum did not generate any insights because the signal was too perturbed. While the quality of the recording allows us to understand the message and hearing the pitch, the pitch tracking algorithms in different analyses software packages (Wavesurfer, Praat) logged values that were unreliable. We would like to point out though that we believe that there is a relationship between the pitch accents in this sequence and that the trigger of downstep can transcend the limits of an intermediate phrase.

Panel 4.1: "My rights as a citizen"

Panel 4.2: "My rights as a citizen."

The final panels for our gestural/intonational microanalysis are panels 4.1 and 4.2. In this segment Mary-Jane has taken advantage of a prop that she brought to the Town Hall meeting to dramatize her point: a hemp-seed chocolate chip cookie in a ziploc bag. She holds up the cookie and addresses the audience, showing them the offending item which is made with hemp seed (the seed of the marijuana plant). She attempts to highlight the absurdity of the government's decree that marijuana is illegal by claiming that she does not have the right to eat a chocolate chip cookie. She ends her performance (one of the issues for us is that we cannot be sure that this is not a rehearsed, prepared speech) with a dramatic flourish, by uttering two parallel constructions, with the same syntactic structure, in a slow and dramatically delivered style. They are presented below, in column format to highlight the parallelism. In conversation analytic transcription conventions, the periods in parenthesis mean that there were significant pauses in the speech stream.

1	2
when did I lose the right	when did I lose my rights
(.)	(.)
to eat	as a citizen of this country
(.)	(.)
the food	to put into my body
(.)	(.)
I want	and listen to whatever
(.)	(.)
to eat.	or watch whatever I want.

In this segment, not only is the syntactic/constructional parallelism evident above, but we further note that there occur to sequences of pitch accents in these

two utterances with downstep quality, despite the fact that in standard Tones and Breaks Indices theory it is thought to be the case that intonational downstep cannot happen across intonational phrase boundaries (Beckman & Ayers 1994). A downstepped H* accent is a type of high star accent, except that the tonal realization is being influenced by the accent preceding the high tone. Often a L+H* accent (a bitonal pitch accent) triggers downstep, so it is thought not to be able to occur across an intermediate phrase, because the trigger is displaced to the preceding intonational phrase. In this case, we would like to claim that for reasons of parallelism and emphasis, downstepped H* (notation: !H*) accents were used in places that have not been previously documented. These !H* accents coincide in our data with the apices of the gestures as well, since Mary-Jane rocks her entire body back and forth approximately 30 degrees to land at the apex of these movements on the !H* accents in both of these utterances.

This brings us to a more general discussion of the correlation between pitch accents and apices.

4.5   Pitch accents & Apices

In order to establish some kind of a measure of what the relation is between the occurrence of an accent and the occurrence of an apex, we counted the number of times Mary-Jane produced accents, that is, prosodic prominences, as well as the number of times where her gesturing included an apex, the peak of the gesture. The numbers are shown in the table below.

Segment	Dur in Sec	Apices	PA	Co-occurring Apices & PA		PA but no Apices	Apices but no PA
1. over	14.23	16 (100%)		16 (100%)			0 (100%)
			24 (100%)	16 (66.7%)		8 (33.3%)	
2. excuse me, hello	11.72	11 (100%)		11 (100%)			0 (100%)
			15 (100%)	11 (73.3%)		4 (26.7%)	
3. tobacco	14.26	16 (100%)		15 (93.8%)			1 (6.3%)
			27 (100%)	15 (55.6%)		12 (44.4%)	
4. lose the right	19.15	28 (100%)		26 (92.8%)			2 (7.1%)
			32 (100%)	26 (81.3%)		6 (18.8%)	
Total	59.36	71 (100%)		68 (95.7%)			3 (4.2%)
			98 (100%)	68 (69.4%)		30 (30.6%)	

Table 1: "(Co-)occurrence of Pitch Accents and Gestural Apices"

The general tendency apparent from this table comparing columns 'Apices' and 'PA' is that there are always more pitch accents than apices in any given segment. Note that we are not comparing the segments with each other, we have merely stated the numbers separately as to make the counting more transparent and easier. As we can see from column 'PA but no Apices', in verbally rendered speech, information is highlighted by acoustic prominences that obviously must not have visually co-occurring gestural apices. On the other hand, it only rarely happens that there is an apex occurring without a pitch accent (see column 'Apices but no PA'). This suggests to us either that some prosodic prominences are just rhythmic in nature or that not all prominences are semantically 'worthy' of being marked by co-speech gestures. It is possible also that the information given on these two different planes is complementary in nature, highlighting *different parts* of the message.

In column, 'Co-occurring Apices & PA', the numbers in the unshaded box (e.g. 16 (100%)) indicates that whenever there is a gestural apex, there also is a pitch accent. The number in the shaded box indicates that only 66.7% of all pitch accents were accompanied by an apex, a beat gesture. In terms of the total

numbers, we find that 95.7% of all apices were accompanied by a pitch accent whereas only 69.4% of all pitch accents were additionally marked by a gestural apex.

We have looked for the co-occurrence of accent location and gestural apex. However, it may be promising also to differentiate between the accent types, which we have not done so far. It is possible that certain (less prominent) accent types such as a downstepped !H* or a L* may be reinforced by gestural means. However we can make no claims regarding this hypothesis.

5   Discussion

We found that speech and gesturing are two different channels/modes of information transfer which allow for different content to be transmitted. If we assume the validity of Bolinger's (1986) claim that "gesture and speech stem from the same semantic intent […]" then we commit ourselves to the notion that some degree of preplanning is involved in generating not only speech output but also gestural output in order to convey information on different planes. How information is structured and divided up across the two channels is not understood at this point. From our data it appears that complimentary and contextual information is transmitted via gestures while concrete assertions are made explicit via speech. We also do not know what constraints exist on (pre-) planning complex gestures that we know are time aligned with linguistic structure in the final output.

Gestures are not just involuntary movements but finely coordinated structures of motion, aligned with semantic content. Since speech can be understood over the telephone or in the dark (with a complete absence of visual cues), we must assume that gestures facilitate the information transfer from the speaker to the listener/viewer but are not necessary for successful transmittal of

content. Gestures play an important role for the naturalness of speech and for cuing speaker stance. On the other hand, it is also known that speakers gesture while speaking when nobody is there to see them. Therefore, it appears that gesturing is not just a facilitation device for the listener/viewer but a mode of self expression for the speaker.

Gesturing exhibits cross-cultural differences, and it is, just as other communicative actions, learned behavior. The mechanisms of acquiring co-speech gesturing consists of coordinating the individual tasks of the complex gestures with each other (for example, lift right arm, rotate palm upward, release arm in this constellation) and then coordinating these motions with speech so that points of informational prominence in speech are accompanied for example by apices in gesture.

An infant as young as a few hours displays the ability to mimic the sticking out of one's tongue when prompted (Meltzoff & Moore 1977, p. 78). This suggests that our cognitive systems provides for learning by example and imitation. The task is a formidable one, as inverse mapping has to take place: the infant observes tongue movement via the visual channel, and then has to map the observed movements to own motor patterns of the articulators (opening lips, lower jaw, extending tongue) in order to perform the task of sticking out the tongue.

It appears that the same type of mechanisms should be involved to learn other motor skills such as finely coordinating hand, arm and body movements to be timed with speech: based on our casual observation, often the offset of a complex gesture co-occurs with the end of a prosodic phrase (intermediate or intonational phrase) or as we have shown in this paper, apices, the peaks of the gestures, co-occur with pitch accents. An interesting test case would be an investigation of the type and timing of co-speech gesturing of people who have been blind since birth.

If the interpretation of intonation contours, that is, if the interpretation of prosodic focus is partly determined by context, we have to take into account that this context is not just provided verbally or relates back to knowledge the interlocutors possess already. Rather, in face-to-face communication the gestural channel is able to provide the speaker's stance, or part of the context in which to interpret the utterance. It appears though that the verbal and gestural channels are interpreted simultaneously and holistically so that the semantic content can be recalled but the presentation and structure of information is more fleeting in nature and thus cannot easily be teased out by the receiver. That is, we tend to remember meaning rather than form but also make inferences which let listeners arrive at an interpretation (Bransford, Barclay, and Franks 1972).

6   Conclusion

In this work we have carried out a case study of the coordination of spontaneous speech with gesture, focusing on intonational alignment with pitch peaks (after Loehr 2004), and have found that wherever there is a gestural apex in our data, there is also a pitch accent. The reverse, however, is not true, because pitch accents often occur without gestures or the apices are phase-shifted from the gestural peaks. Our results concur with those of Loehr (2004) for laboratory data, and indicate that gestural phenomena are in robust co-occurrence with pitch accents in both laboratory and spontaneous speech. Our findings also include the discovery of a downstep relationship across intonational phrase boundaries, as well as the pervasive use of lexical, syntactic, gestural, and intonational parallelism in the performance of a speaker under high pressure in an affect-laden, spontaneous communicative situation. While the intonational and gestural alignments were observed within and across intonational phrases, metaphorical gestural spatialization had a larger domain, it took longer to

unfold, and sketched out complementary information on this speaker's notion of the relationship between a political representative, the government, and outside entities.

7   Appendix: Transcription of Town Hall Meeting

Town Hall meeting at St. Cyril's, midtown Tucson school, February 2001. Filmed with Sony DVTR8 mini-DV camera, sound aligned with audience microphone, quicktime clip name: kolbestcyril2(2/18,DVTR8).mov

Kolbe   and since you wanna comment on [this

            [I'm sorry]

            We're gonna] get your comment[

 M-J   ]my--

            my only problem with--

(0:04)   n- n- n- not legalizing all drugs?

Green   right.

M-J   you take the criminal element out of it,

            when you end prohibition,

            (.)

            this is what we saw in the thirties.

            the drug war will be over.

(0:20)   we won't have to ship guns and military to Colombia,

            it will be over.

            Bush won't have to talk to Fox about the drug war,

            it will be

            (.)

            over.

Aud   [xxclappingxx]

           [there will not BE a drug war]

           if you legalize drugs.

Kolbe   ok [[calm down,

(0:30)   calm down,

           (.)

           calm down,]]

 M-J   [[you are making it a criminal]] enterprise,

Aud: boooo]

           [you have got-]

Aud: [boooo]

           the government has made it a criminal enterprise,

           the government is making money,

           you are making money hand over fist,

(0:40)   you are building prisons so fast it's disgusting,

           you're not putting any of that money into education,

           you're locking those-

           you're building prisons to lock those kids up that you're talking

           about today,

           they're going to go to prison.

(0:50)   because they're going to experiment with drugs,

           like every one has done

           Since day one.

Kolbe   ok,

           well I--[

M-J   ]we experiment.

           alcohol and tobacco kill people.

           marijuana doesn't kill anybody.

excuse me,

hello-o:

(.)

Kolbe   ok.

      (.)

      well--[

M-J   ]why aren't you talking about alcohol and cigarettes?

what you've done is you've taken that industry,

and you are putting it on third world.

Kolbe   what--

M-J   they are consuming the tobacco now.

that's where all the tobacco money is coming from.

Kolbe   [we're spending]

M-J   [they're still making] a profit on cigarettes cause they're selling

them to all the [people]

that we're dumping our chemicals on

Kolbe   [we're--]

we're spending a lot of money on a--

(.)

on education,

which is what I think—

(.)

on tobacco,

which is what I think we need to be doing,

(.)

[spending a lot on that].

M-J   [proper name] wants to make this

a schedule one narcotic.

A chocolate chip cookie,

ok?

a chocolate chip cookie,

this is what they tried to run through on October the thirtieth,

they want to make

(.)

hemp seed

a schedule one narcotic.

when did I lose the right,

(.)

to eat,

(.)

the food,

(.)

I want,

(.)

to eat.

(1:40)   when did I lose my rights

(.)

as a citizen of this country

(.)

to put in my body

(.)

and listen to whatever

(.)

or watch whatever I want.

Kolbe   well,

(.)

uh,

there--

there always have been limits on doing--

on some things.

you do—

we do--

you do not have[

M-J   ]I limit myself,

[you don't limit me thank you].

Kolbe   [y:--

y:--

you] don't have absolute rights to do everything.

M-J   yes we do.

Kolbe   and we do--

(.)

we never have had.

References

Beattie, G., & Shovelton, H. (1999) An Experimental Investigation of the Role of Iconic Gestures in Lexical Access Using the Tip of the Tongue Phenomenon. *British Journal of Social Psychology*, 90, pp. 35-56.

Beckman, M. E. & Ayers, G. E. (1994) *Guidelines to ToBI Labelling (version 2.0)*. The Ohio State University.

Berthoz, A. (1997) Le Sens du Mouvement. Paris: Editions Odile Jacob.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972) Sentence Memory: a Constructive Versus Interpretive Approach. *Cognitive Psychology* 3, pp. 193-209.

Bohman, J. (1996) *Public Deliberation.* Cambridge, MA: MIT Press.

Bollinger, D. D. (1986) Intonation and Gesture. *American Speech*, 58.2, pp. 156-174.

Brown, P., & Levinson, S. C. (1993) "Uphill" and "Downhill" in Tzeltal. *Journal of Linguistic Anthropology* 3(1), pp. 46-74.

Browman, C. P., & Goldstein, L. (1990) Tiers in articulatory phonology, with some implications for casual speech. In T. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, pp. 341-376. Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992) Articulatory Phonology: An overview. *Phonetica,* 49, pp. 155-180.

Cassell, J., McNeill, D., & McCollough, K. E. (1999) Speech-gesture mismatches: Evidence for One Underlying Representation of Linguistic and Nonlinguistic Information. *Pragmatics and Cognition*, 7(1), pp. 1-33.

Cassell, J. (2000) A framework for Gesture Generation and Interpretation. In R. Cipolla and A. Pentland (eds.) Computer Vision in Human-Machine Interaction. Cambridge: Cambridge University Press. http://www.soc.northwestern.edu/justine/publications/gesture_wkshop.pdf

Chomsky, N. (2000) Minimalist Inquiries: The Framework. In *Step by Step*, R. Martin, D. Michaels, and J. Uriagereka (eds.), pp. 89-155. Cambridge, MA: MIT Press.

Cosnier, J. (1982) Communications et langages gestuels. In J. Cosnier, A. Berrendonner, J. Coulon, and C. Orecchioni (eds.) *Les voies du langage: Communications verbales, gestuelles, et animales*, pp. 255-304. Paris: Dunod.

Enfield, N. (2003) Producing and Editing Diagrams Using Co-Speech Gesture: Spatializing Non-spatial Relations in Explanations of Kinship in Laos. *Journal of Linguistic Anthropology* 13, pp. 7-50.

Ekman, P. & W. Friesen (1972) Hand Movements. *The Journal of Communication* 22, pp. 353-374.

Farnell, B. (1995) Human Action Signs in Cultural Context: the Visible and the Invisible in Movement and Dance. Metuchen, NJ: Scarecrow Press

Fadiga, L. & Craighero, L. (2003) New Insights on Sensorimotor Integration: From hand action to speech perception. *Brain and Cognition* 53, pp. 514-534.

Fodor, J. (1983) *The Modularity of Mind.* Cambridge, MA: MIT Press.

Halle, M. & Stevens, K. (1959) Analysis by Synthesis. In W. Wathen-Dunn and L.E. Woods (eds.) *Proceedings of the Seminar on Speech Compression and Processing.* AFCRC-TR-59-198, December 1959, Vol. II, paper D7.

Harper, L., Loehr, D. & Bigbee, A. (2000) Gesture is not just Pointing. Presented at and published by the International Conference on Natural Language Generation.

Haviland, J. B. (1993) Anchoring, Iconicity, and Orientation in Guugu Yimidhirr Pointing Gestures. *Journal of Linguistic Anthropology*, Vol. III(1), pp. 3-45.

Hirschberg, J. & Ward, G. (1992) The Influence of Pitch Range, Duration, Amplitude and Spectral Features on the Interpretation of the rise-fall-rise Intonation Contour in English. *Journal of Phonetics*, 20, pp. 241-251.

Goodwin, C. (1981) *Conversational Organization. Interaction between Speakers and Hearers.* New York: Academic Press.

Kelly, S., Barr, D., Church, R., & Lynch, K. (1999) Offering a Hand to Pragmatic Understanding: The Role of Speech and Gesture in Comprehension and Memory. *Journal of Memory and Language,* 40, pp. 577-592.

Kendon, A. (1972) Some Relationships between Body Motion and Speech. In Siegman, A. & Pope, B., (eds.) *Studies in Dyadic Communication*, pp. 177-210. New York: Pergamon Press.

Kendon, A. (1980) Features of the Structural Analysis of Human Communicational Behaviour. In von Raffler-Engel, W. (ed.), *Apsects of Nonverbal Communication*. Swets & Zeitlinger, Lisse, Netherlands, pp. 29.43.

Kendon, A. (1996) Gesture in Language Acquisition. *Multilingua* 15, pp. 201-214

Krifka, M. (2005) A Functional Similarity between Bimanual Coordination and Topic/Comment Structure. Paper presented at the Blankensee Colloquium,

Berlin Schmöckwitz: Language Evolution: Cognitive and Cultural Factors, July 14-16, 2005.

Liberman, P. & Blumstein, S. (1990) Speech Physiology, Speech Perception, and Acoustic Phonetics. *Cambridge Studies in Speech Science and Communication*. Cambridge: Cambridge University Press.

Liberman, A. & Mattingly, I. (1985) The Motor Theory Revised. *Cognition* 21, pp. 1-36.

Liberman, M. & Pierrehumbert, J. (1984) Intonational Invariants under Changes in Pitch Range and Length. In M. Aronoff & R. Oehrle (eds.) Language Sound Structure. MIT Press, Cambridge, MA.

Loehr, D. (2004) *Gesture and Intonation.* Doctoral Dissertation, Georgetown University, Washington, DC.

Marantz, A. (to appear) Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review.* Accessible from author's website: http://web.mit.edu/linguistics/www/marantz/index.html

McNeill, D. (1992) *Hand and mind.* The University of Chicago Press.

McNeill, D. (2000) *Language and gesture.* Cambridge: Cambridge University Press.

McNeill, D., Quek, F., McCullough, K.-E., Duncan, S., Furuyama, N., Bryll, R., Ma, X.-F. & Ansari, R. (2001) Catchments, Prosody and Discourse. *Gesture* 1:1, pp. 9-33.

Meltzoff, A. N., & Moore, M. K. (1977) Imitation of Facial and Manual Gestures by Human Neonates. Science, 198, pp. 75-78.

Pierrehumbert, J. (1980) *The Phonology and Phonetics of English Intonation.* Doctoral Dissertation, MIT.

Pierrehumbert, J. & Hirschberg, J. (1990) "The Meaning of Intonation Contours in the Interpretation of Discourse." In Cohen, Morgan, and Pollack, (eds.) *Plans and Intentions in Communication and Discourse.* MIT Press, pp. 271-311.

Rimé, B. (1982) The Elimination of Visible Behavior from Social Interactions: Effects on Verbal, Nonverbal, and Interpersonal Variables. *European Journal of Social Psychology* 12, pp. 113-29.