

1st International Conference of SFB632:

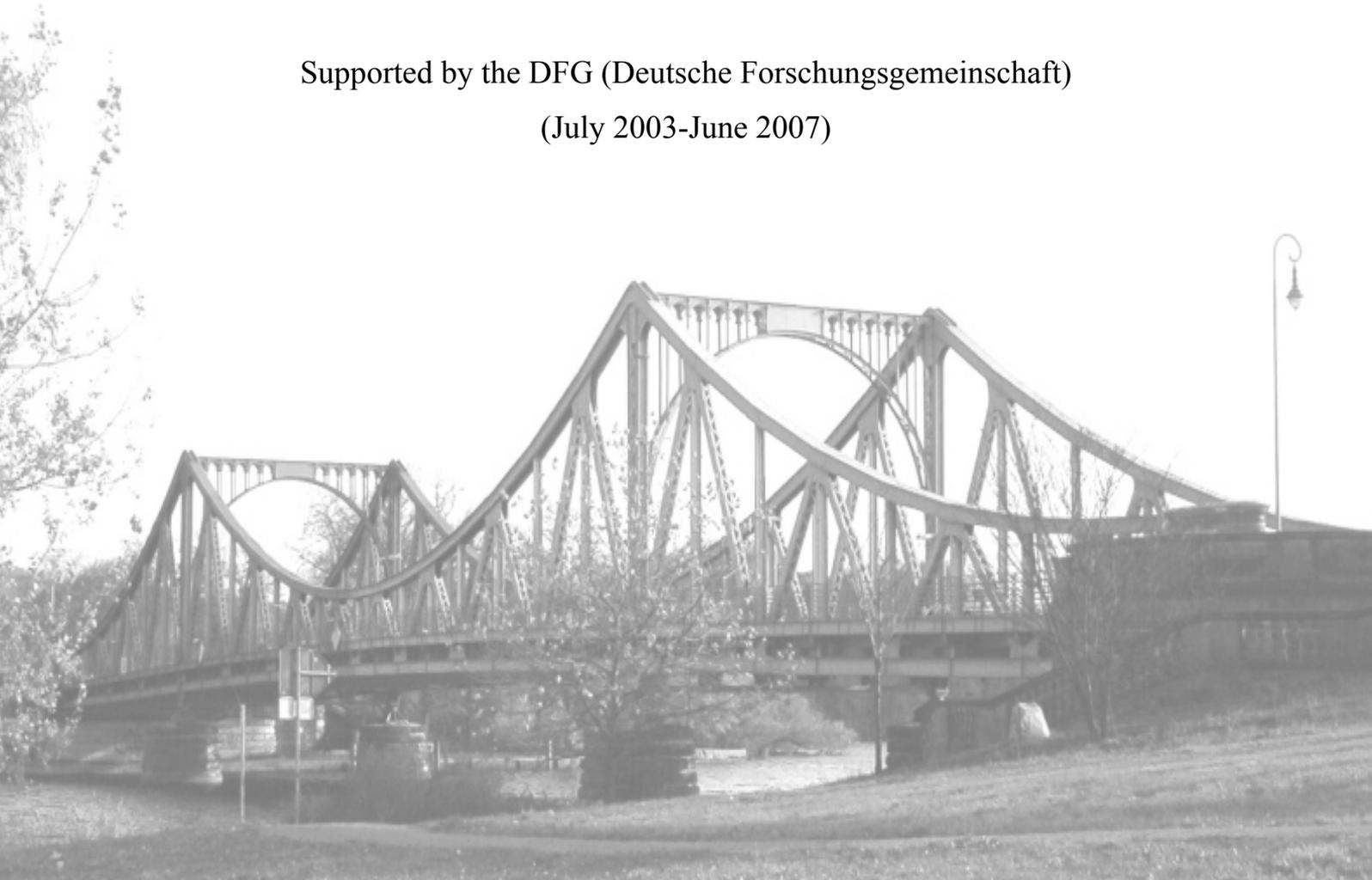
**Information structure between linguistic theory and
empirical methods**

University of Potsdam

Humboldt University of Berlin

June 6-8, 2006, Potsdam

Supported by the DFG (Deutsche Forschungsgemeinschaft)
(July 2003-June 2007)



1st International Conference of SFB632:

**Information structure between linguistic theory and
empirical methods**

University of Potsdam

Humboldt University of Berlin

June 6-8, 2006, Potsdam

<http://www.sfb632.uni-potsdam.de>

Campus I “Neues Palais”, Building 9, Room 1.12

Tuesday, June 6

A Projects: Theory development

12:45-13:00	Coffee
13:00-13:10	Opening
13:10-13:50	Contrastive topics <i>S. Tomioka (University of Delaware)</i>
13:50-14:30	How information structure shapes prosody <i>C. Féry and S. Ishihara (University of Potsdam, Project A1)</i>
14:30-14:45	Coffee break
14:45-15:25	Discourse semantics and prosody, A case study: The interpretation of tag questions <i>N. Asher and B. Reese (University of Texas at Austin)</i>
15:25-16:05	Quantification and topicality <i>C. Endriss and S. Hinterwimmer (University of Potsdam, Humboldt University of Berlin, Project A2)</i>
16: 05-16:20	Coffee break
16:20-17:00	S-focus and relativized stress F <i>M. Rooth (Cornell University)</i>
17:00-17:40	Scalar implicatures – which theory best explains their processing and acquisition <i>T. Reinhart (Tel Aviv University)</i>

Wednesday, June 7

D Projects: Data management

8:50-9:00	Coffee
9:00-9:40	A lexical perspective on discourse structure and semantics <i>B. Webber (University of Edinburgh)</i>
9:40-10:20	From data to insights: exploiting linguistic data <i>S. Dipper, M. Götze, J. Ritz and S. Skopeteas</i> <i>(University of Potsdam, Project D1 and D2)</i>
10:20-10:35	Coffee break
10:35-11:15	Why negative adverbials of degree, manner, and frequency must be focussed <i>K. E. Kiss</i> <i>(Linguistic Institute of the Hungarian Academy of sciences)</i>
11:15-11:55	Towards a typology of information structure: insights from cross-linguistic data collection <i>S. Hellmuth and S. Skopeteas (University of Potsdam, Project D2)</i>
11:55-13:30	Lunch

B Projects: Language comparison and language change

13:30-14:10	Information structure and OV order <i>T. Hróarsdóttir (University of Tromsø)</i>
14:10-14:50	The role of information structure in word order variation and word order change <i>R. Hinterhölzl (Humboldt University of Berlin, Project B4)</i>
14:50-15:05	Coffee break
15:05-15:45	Is there a focus position in Aghem? <i>L. Hyman and M. Polinsky</i> <i>(University of California Berkeley, University of California San Diego)</i>
15:45-16:25	Subject focus in West African languages <i>I. Fiedler, K. Hartmann, B. Reineke, A. Schwarz, and M. Zimmermann (Humboldt University of Berlin, Project B1 and B2)</i>
16:25-16:40	Coffee break
16:40-17:20	Been there- marked that: A theory of second occurrence focus <i>D. Büring (UCLA)</i>
19:00	Dinner

Thursday, June 8

C Projects: Cognition and information structure

8:50-9:00	Coffee
9:00-9:40	Crosslinguistic studies on the role of information structure for ambiguity resolution <i>B. Hemforth (Université de Provence)</i>
9:40-10:20	Information structure and the anticipation of discourse referents <i>T. Weskott, R. Hörnig, *E. Kaiser, C. Féry, G. Fanselow, S. Kern, and R. Kliegel</i> <i>(University of Potsdam, * University of Southern California, Project C1)</i>
10:20-10:35	Coffee Break
10:35-11:15	Children's association with focus <i>Kenneth F. Drozd (University of Aarhus)</i>
11:15-11:55	The acquisition of the additive focus particle <i>auch</i> in German: A matter of competence or performance? <i>F. Berger, A. Müller, B. Höhle, M. Schmitz, and J. Weissenborn</i> <i>(University of Potsdam, Humboldt University of Berlin, Project C3)</i>
11:55-13:00	Lunch
13:00-15:00	Poster Session 32 Posters of SFB632

Preface

Information structure is intrinsically multimodular and can only be successfully addressed when several aspects of grammar are considered at once. On the theoretical side, semantic models need input from phonology and syntax; similarly, both phonology and syntax show the most interesting results when supported by other grammatical components. On the empirical side, we are now going through an area of experimental investigation which promises to be highly profitable for the research on information structure, shedding light in particular on the aspects of typology, psycholinguistics and language acquisition.

Invited Speakers are:

- Nicholas Asher (University of Texas at Austin)
- Daniel Büring (UCLA)
- Kenneth F. Drozd (University of Aarhus)
- Barbara Hemforth (University de Provence)
- Thorbjörg Hróarsdóttir (University of Tromsø)
- Larry Hyman (University of California Berkeley)
- Katalin E. Kiss (Hungarian Academy of sciences)
- Maria Polinsky (University of California Berkeley)
- Tania Reinhart (Tel Aviv University)
- Mats Rooth (Cornell University)
- Satoshi Tomioka (University of Delaware)
- Bonnie Webber (University of Edinburgh)

SFB632 present themselves with 8 talks and 30 posters.

Content

Invited Speakers

Asher, N. and B. Reese: Discourse Semantics and Prosody, A Case Study: The Interpretation of Tag Questions	1
Büring, D.: Been There — Marked That: A Theory of Second Occurrence Focus.....	5
Drozd, K.: Children's Association with Focus	9
Hemforth, B.: Crosslinguistic Studies on the Role of Information Structure for Ambiguity Resolution	17
Hróarsdóttir, T.: Information Structure and OV Order	23
Hyman, L. and M. Polinsky: Is There a Focus Position in Aghem?	27
E. Kiss, K.: Why Negative Adverbials of Degree, Manner, and Frequency Must be Focussed	31
Reinhart, T.: Scalar Implicatures – Which Theory Best Explains Their Processing and Acquisition?	37
Rooth, M.: S-focus and Relativized Stress F	43
Tomioka, S.: Contrastive Topics	47
Webber, B.: A Lexical Perspective on Discourse Structure and Semantics	51

SFB632: Information Structure

Project A1: Phonology and syntax of focusing and topicalisation

Féry, C. and S. Ishihara: How Information Structure Shapes Prosody	55
Fanselow, G.: Autonomy of Syntax and Information Structure	59

Project A2: Quantification and information structure

Endriss, C. and S. Hinterwimmer: Quantification and Topicality	63
Endriss, C. and S. Hinterwimmer: Interpreting Topics in Quantificational Structures	69

Project A3: Rhetorical structure in spoken language: modeling of global prosodic parameters

Jasinskaja, E., Kölsch, U. and J. Mayer: Anaphora Resolution in Discourse: The Role of Pitch Range and Pause75

Jasinskaja, E.: Pragmatics and Prosody of Implicit Discourse Relations: The Case of Restatement79

Project A4: Focus evaluation, anaphoricity, discourse coherence

Kasimir, E., Şahin, Ö. and J. Böhme: Beyond Topic and Focus: Extraposition in Turkish.....83

Project B1: Focus in Gur and Kwa languages

Project B2: Focusing in Chadic languages

Project B4: The role of information structure in the development of word order regularities in Germanic

Fiedler, I. et.al.: Subject Focus in West African Languages.....89

Hartmann, K. and M. Zimmermann: Focus in Chadic Languages93

Hinterhölzl, R.: The Role of Information Structure in Word Order Variation and Word order Change.....97

Fiedler, I., Reineke, B. and A. Schwarz: Focus in Gur and Kwa 101

Petrova, S. and A. Schwarz: Discourse Structure and Information Packaging in Cross-Linguistic Perspective 107

Donhauser, K. et al.: Information Structure and Word Order in the Early Germanic Languages and its Analysis in a Linguistic Database 113

Project C1: Contextually licensed non-canonical word order in language comprehension

Weskott, T. et al.: Information Structure and the Anticipation of Discourse Referents119

Hörnig, R. et al.:Contextual Licensing of Marked Word Order in Spatial Descriptions123

Hörnig, R. et al.: Structural Markedness and Structural Facilitation in Comprehending Descriptions of Unknown and Known Spatial Layouts..... 127

Project C3: The acquisition of focus marking in unimpaired and impaired first language acquisition: prosodic, syntactic, and lexical aspects

Berger, F. et al.: The Acquisition of the Additive Focus Particle auch in German: A Matter of Competence or Performance? 131

Schmitz, M. et al.: Recognition of the Prosodic Focus position in German-learning Infants..... 137

Project C4: Prosody and information structure as forms of 'input' in second language acquisition

Sennema, A., van de Vijver, R. and A. Zimmer-Stahl: The Influence of Focus Markers on Word Processing and Word Recall in a Second Language..... 143

van de Vijver, R., Sennema, A. and A. Zimmer-Stahl: An Analysis of Pitch and Duration of Experimental Material 149

Vasishth, S.: Processing Information Structure 153

Project D1: Linguistic database for information structure: annotation and retrieval

Dipper, S. et al.: From Data to Insights: Exploiting Linguistic Data..... 157

Chiarcos, C.: Sustainability of Linguistic Data 161

Stede, M. et al: ANNIS: A Linguistic Database for Complex Multilevel Annotation.. 167

Project D2: Typology of information structure

Hellmuth, S. and S. Skopeteas: Towards a Typology of Information Structure: Insights From Cross-linguistic Data Collection..... 171

Skopeteas, S. et al.: Information Structure in the Project Languages of D2 'Typology of Information Structure' 175

Wang, B. and C. Féry: Split-noun Sentences in Mandarin and in German: Production and Perception 181

Project D3: Signal parameters connected to prominence and phrasing within spoken utterances in different languages

Pompino-Marschall, B.: Tone, Intrinsic f0, and Intonational Focus Marking in Chichewa 187

Jannedy, S.: Prosodic Edge Marking in Ewe 191

Discourse Semantics and Prosody, A Case Study: The Interpretation of Tag Questions

Nicholas Asher, Brian Reese

The University of Texas at Austin

bjreese@mail.utexas.edu nasher@mail.utexas.edu

This talk begins with a general discussion of how prosody is a crucial bit of information in determining the discourse structure and content of a discourse. We will review some of the basic findings by Pierrehumbert and Hirschberg (1990), Hirschberg and Ward (1992), Gunlogson (2003) and others before focusing on particular types of complex speech acts that are performed with intonation. Our case study involves reversed-polarity tag questions, i.e., a declarative *anchor* in a paratactic or supplemental relation with an interrogative *tag* (Huddleston and Pullum, 2002). We do not address constant-polarity tag questions and tag questions containing an imperative, interrogative or exclamative anchor. Illustrative examples are given in 1 and 2¹.

- (1) a. But I don't have time to rewrite it, do I?
b. You've really haven't been fed, have you.
- (2) a. All of us have been afraid of something, haven't we?
b. That sounds like it was right in front of my house,
doesn't it.

More importantly, (1a) and (2a) differ from (1b) and (2b) prosodically. These prosodic differences are correlated with semantic and pragmatic differences. The (a)-examples, for instance, express more uncertainty or doubt

¹ These examples are from Part-II of the Santa Barbara Corpus of Spoken American English. In all, we extracted 220tokens of tag questions from the corpus. Naturally occurring examples are indicated with typewriter font.

on the part of the speaker toward the truth of the anchor than do the (b)-examples (Huddleston and Pullum 2002, Sadock 1974, Ladd 1981). The correct description of the prosodic difference, however, is controversial. The most common characterization is that the (a)-examples have a rising pitch contour on the tag, while the (b)-examples have falling pitch (Huddleston and Pullum 2002, Sadock 1974). (Falling/rising intonation is indicated orthographically with a period/question mark.)

But are others for whom prosodic differences having to do with accent and phrasing are key (Sadock 1974, Ladd 1981). Ladd (Ladd 1981) is the strongest advocate of this position, casting aside the falling vs. rising dichotomy entirely, the pivotal distinction being between tags that contain a nuclear accent vs. those that do not. Nuclear tag questions consist of two intonational phrases encompassing the anchor and the tag and each of which contains a separate nuclear pitch accent. Postnuclear tag questions contain only one intonational phrase and no separate nuclear pitch accent on the tag. The nuclear/postnuclear distinction, according to Ladd, corresponds semantically to the falling/rising distinction. Ladd also notes that the distinction appears to correlate with differences in the scope of negation, in addition to other “unexpected meaning shifts” (Ladd 1980).

We explore these distinctions and develop an analysis of tag questions that aligns their prosodic and intonational properties with their discourse function. Our formal framework is Segmented Discourse epresentation Theory (SDRT), a dynamic semantic framework for computing the rhetorical connections between utterances in discourse and dialogue (Asher and Lascarides, 2003).

References

- Asher, Nicholas, and Alex Lascarides (2003). *Logics of Conversation*. Cambridge University Press.
- Büring, Daniel, and Christine Gunlogson. 2000. Arent positive and negative polar questions the same? UCSC/UCLA.
- Gunlogson, Christine. 2003. True to form: Rising and falling declaratives as questions in English. *Outstanding Dissertations in Linguistics*. New York: Routledge.
- Hamblin, C. L. 1958. Questions. *Australasian Journal of Philosophy* 36:159-168.
- Han, Chunghye. 2002. Interpreting interrogatives as rhetorical questions. *Lingua* 112:201-229.
- Hirschberg, Julia, and Gregory Ward. 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* 20:241-251.
- Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge University Press.
- Ladd, D. Robert. 1980. *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.
- Ladd, D. Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In *Papers from the 17th Regional Meeting of the Chicago Linguistic Society*, 164-171. University of Chicago.
- Lieberman, Mark, and Ivan Sag. 1974. Prosodic form and discourse function. In *Papers from the 10th Regional Meeting of the Chicago Linguistic Society*, 416-427. University of Chicago.
- Marandin, Jean-Marie, Claire Beyssade, Elisabeth Delais-Roussarie, Annie Riolland, and Michel de Fornel. 2005. The meaning of final contours in French. Submitted for publication.
- Pierrehumbert, Janet, and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication*, ed. Philip Cohen, Jerry Morgan, and Martha Pollock, 271-311. Cambridge, MA: MIT Press.

Sadock, Jerrold M. 1974. *Towards a Linguistic Theory of Speech Acts*. New York, NY: Academic Press.

Safarova, Marie. 2005. The semantics of rising intonation in interrogatives and declaratives. In *Proceedings of SuB9*, ed. Emar Maier, Corien Bary, and Janneke Huitink, 355369. www.ru.nl/ncs/sub9

Been There — Marked That: A Theory of Second Occurrence Focus

Daniel Büring

UCLA

buring@humnet.ucla.edu

In this talk I present a theory of focus realization in English, i.e. a set of rules and principles that associates syntactic structures with F(ocus)-marking in them with prosodic structures including pitch accents. The theory is a blend of existing theories as presented in Selkirk (1995, 2005), Truckenbrodt (1995, 2006, in press), Samek-Lodovici (2002), Féry and Samek-Lodovici (2006), Büring and Gutiérrez-Bravo (2001) and Büring (2001a,b), but covers new empirical ground in that it *captures second occurrence foci* (2OFi).

The phenomenon of 2OF is illustrated in (1) (from Hajičová et al. (1998)):

- (1) (Everyone already knew that Mary only eats vegetables.)
If even PAUL_F knew that Mary only eats vegetables_F, then he should have suggested a different restaurant

The final occurrence of *Vegetables* in (1) is a 2OF; it is a focus since it associates with *only*, yet it is not pitch accented, as foci normally are. 2OF has received considerable attention in the literature (e.g. Gussenhoven (1984), Hajičová(1973, 1984), Partee (1991, 1999)). As perhaps pointed out most clearly in Partee (1999), 2OF examples seem to force us to question either the assumption that foci always receive a prosodic realization, or the assumption that particles like *only* really (directly) associate with focus.

Beaver and Clark (2003), however, makes a convincing argument that the associates of *only* and the likes *are* grammatical foci. At the same time, it has

been reported that 2OFi *are* prosodically marked, albeit not (always) by pitch accent, as is the custom for well-behaved ‘normal’ foci, but by lengthening (Bartels (2004), Rooth (1996, 2004), and, to an extent, Krifka (2004)); their reports have been confirmed by recent, more systematic studies reported in Beaver et al. (2004), Féry and Ishihara (2005), and Jaeger (2004).

I assume that the lengthening of foci is the consequence of metrical *stress*, i.e. prominence in the metrical structure. So all foci, ordinary and 2O, are made prominent by stress. In a nutshell:

- Second occurrence foci are not marked by pitch accent when they occur in post-nuclear position.
- They are, however, marked by lengthening. in post-nuclear position.
- 2OFi are marked by pitch accent when pre-nuclear.

Unfortunately, these generalizations, to a certain extent, beg the main question: How come the 2OF does not get the nuclear accent? What does the element with the nuclear accent — the ‘ordinary’ or ‘first occurrence focus’ (1OF), as it were — have, that the element in 2OF does not? Unless we answer that question, the first bullet point above simply restates the definition of nuclear accent. In short, we need a definition of 2OF *vis-a-vis* 1OF that does not itself make reference to their prosodic realization. Once we have that, and only then, can we try to work the above generalizations into a theoretical account.

I show that a view implicit in most of the existing literature, viz. that 2OFi are simply *given* foci, is not tenable. I suggest that the crucial distinction is between foci whose *domain* is *maximal*, i.e. the entire sentence, and foci whose domain is smaller. A focus with a non-maximal domain will potentially be realized as a 2OF. What complicates matters a bit is that the distinction between foci with maximal v. non-maximal domains cuts across the Given/New

distinction, and that some foci have more than one domain; so clarifying the proposal requires a close look at focus representation and interpretation, and requires the use of indexed foci, as done e.g. in Jacobs (1991), Kratzer (1991) and Wold (1996), among others. Foci and focus domains are determined semantically/pragmatically, using a mix of the theories in Rooth (1992) and Schwarzschild (1999).

The second challenge is to explicate why the 1/2OF (or maximal/nonmaximal domain F) distinction yields the prosodic consequences it does: Stress for all foci, pitch accents for all 1OF_i, but for 2OF only if they precede a 1OF. Here, the account I offer involves two steps.

First, perhaps obviously, **stress, rather than accent, is the basic realization of focus**: the metrical structure of a sentence will be constructed in such a way that focussed elements are metrically strong, i.e. receive phrase level or sentence level stress (see previous references). Pitch accent assignment then proceeds ‘mechanically’, as it were, given the prosodic structure; it does not ‘see’ features like F(ocus), or any aspect of syntactic structure for that matter.

The second step is that **foci with a non-maximal domain are banned from bearing sentence-level stress** (and thus indirectly from bearing the NPA). So if the last metrically strong element in a sentence is a 2OF, the NPA will ‘shift’ to the left, leaving the 2OF stressed (lengthened), but unaccented. If in pre-nuclear position, the 2OF will receive a (non-nuclear) pitch accent, just like any other metrically strong element. Focus Prominence, properly relativized to focus domains, correctly predicts this.

The theory is not only the first, to my knowledge, to capture the basic cases of 2OF, it also predicts a number of new facts regarding the possible configurations of ordinary and 2O foci.

Towards the end of the talk, I address a number of open questions. I tentatively suggest that the domain of a focus should be defined in prosodic, rather than syntactic terms. I also discuss the possibility of finding 2OFi that are not associated with focus sensitive particles such as *only*. And I spend some time discussing the interaction of prosodic structure and focus projection.

References

- Beaver, David, Brady Clark, Edward Flemming, Florian Jaeger, and Maria Wolters. 2004. When semantics meets phonetics: Acoustical studies of second occurrence focus. Ms. Stanford University .
- Féry, Caroline, Shinichiro Ishihara. 2005. Occurrence Focus. Ms. Potsdam University .
- Rooth, Mats. 1996. On the Interface Principles for Intonational Focus. Proceedings of Semantics and Linguistic Theory (SALT) VI, edited by Teresa Galloway and Justin Spence, 202–226. Ithaca, NY: Cornell University.
- Truckenbrodt, Hubert. 1995. Phonological Phrases: Their Relation to Syntax, Focus, and Prominence. Ph.D. thesis, MIT. Published 1999 by MITWPL.

Children's Association with Focus

Ken Drozd

University of Aarhus

engkd@hum.au.dk.

1 Children's association strategies

Experimental studies suggest that children and adults analyze sentences with *only* (*only* sentences) in strikingly different ways (Paterson *et al.* 2003; Crain, Ni, & Conway 1994; Crain, Philip, Drozd, Roeper, & Matsuoka 1992). *Only* contributes to the interpretation of a sentence like (1) or (2) by specifying that the proposition expressed by the sentence holds of the extension of the focus-marked expression in the operator's syntactic scope (*the bird* and *a flag*, respectively) to the exclusion of alternatives to focus in the discourse model, giving rise to the familiar association-with-focus (**AF**) phenomenon. However, whereas adults assign subject-focus interpretations to (1) and object-focus interpretations to (2), previous studies have shown that many children assign a subject focus interpretation to both sentences (the ASSOCIATE-WITH-SUBJECT (**AS**) strategy) while others assign an object focus interpretation to both sentences (the ASSOCIATE-WITH-OBJECT (**AO**) strategy), regardless of the syntactic position of *only*.

- (1) **Pre-Subject *only* Sentence:** Only [the bird]_F is holding a flag.
Subject-Focus Interpretation: The only thing holding a flag is a bird
- (2) **Pre-Object *only* Sentence:** The bird is holding only [a flag]_F .
Object-Focus Interpretation: The only thing the bird is holding is a flag

2 Experiment

Based on previous research on children's performance with universal quantification (e.g., Drozd & van Loosbroek 2006), we hypothesized that children consistently associate *only* with constituents other than the intended focus of a sentence because they do not represent the intended focus as relevant and/or contrastive in the discourse model. We conjectured that children's performance with **AF** should dramatically improve if the intended focus of a sentence were made explicit by means of contrastive intonation or contrastive reference. To test this hypothesis, 216 native Dutch-speaking children and adults were instructed to judge *alleen* sentences like (3a,4a) and (3b,4b) as true or false with respect to pictured contexts like Contexts 1 and 2, respectively.

<p><u>Visual Context 1</u> A boy and a girl are each holding a different brush.</p>	<p><u>Visual Context 2</u> A boy and a girl are each holding a different balloon. The girl is also holding a pie.</p>
<p>(3a) <u>Pre-Object <i>Alleen</i> Sentence 1</u> Het meisje heeft alleen een kwast (Ja) 'The girl has only a brush (Yes)'</p>	<p>(3b) <u>Pre-Object <i>Alleen</i> Sentence 2</u> Het meisje heeft alleen een taart. (Nee) 'The girl has only a pie (No)'</p>
<p>(4a) <u>Pre-Subject <i>Alleen</i> Sentence 1</u> Alleen het meisje heeft een kwast (Nee) 'Only the girl has a brush.(No)'</p>	<p>(4b) <u>Pre-Subject <i>Alleen</i> Sentence 2</u> Alleen het meisje heeft een taart. (Ja) 'Only the girl has a pie (Yes)'</p>

Each participant received 3 trials of each of the 4 visual context-sentence combinations above but in only one of 3 different discourse conditions. In the **NEUTRAL DISCOURSE CONDITION**, test sentences were presented with neutral intonation and in neutral (noncontrastive) referential and discourse contexts. The **CONTRASTIVE INTONATION CONDITION** differed from the Neutral Discourse condition only in that test sentences were presented with contrastive stress on the intended focus of the sentence (e.g., *The girl is holding only a BRUSH*). The

CONTRASTIVE REFERENCE CONDITION differed from the Contrastive Intonation condition only in that the experimenter explicitly mentioned a set of focus alternatives before presenting a test sentence (e.g., *If you want to paint the fence, you need a brush and, of course, a can of blue paint. . . . The girl is holding only a BRUSH*).

3 Results

Participants overall correctly judged Pre-Subject Alleen Sentences app. 85% of the time across all three discourse conditions, but performed poorly with Pre-Object Alleen sentences in all but the Contrastive Reference condition. In the Neutral Discourse condition, participants overall correctly responded to (3a) and (3b) only app. 25% of the time. Further analysis revealed that most of the children and one-third of the adults adopted an AS strategy given this condition. Children's performance significantly improved with Pre-Object Alleen sentences to 60% correct on the Contrastive Intonation condition.. However, 21% of the children adopted an AO strategy on this condition, compared to only 6% given the Neutral Discourse Condition. Children's performance with Pre-Object Alleen sentences significantly improved again on the Contrastive Reference condition. Yet half of the children adopted an AS strategy on this condition as well.

4 Analysis: the AS strategy

I will argue that participants who adopt the **AS** strategy consistently associate *only* with the subject because they represent the denotations of the subjects of our test sentences as relevant and contrastive in the discourse model and mark the subjects of both pre-subject *only* and pre-object *only* sentences as contrastive foci to establish discourse congruence between test sentence and utterance

context. An example analysis extending von Fintel's (1994) association-by-anaphora account of **AwF** is given in (5).

(5) Associate-With-Subject Strategy

- (5a) Discourse Topic: Which child has a brush?
 $[[Q]]^o = \{\mathbf{have}(x, \mathbf{brush}) \mid x \in E \wedge \mathbf{child}(x)\}$
- (5b) Test Sentences: Only_C [the girl]_F has a brush
- (5c) [The girl]_F has only_C a brush
- (5d) Truth Conditions: $\forall p[p \in C \wedge \checkmark p \rightarrow p = \mathbf{have}(\mathbf{girl}, \mathbf{brush})]$
- (5e) Domain Selection: $C = [[Q]]^o$
- (5f) Presupposition: $[[Q]]^o \subseteq \{\mathbf{have}(x, \mathbf{brush}) \mid x \in E\}$ which minimally includes $\mathbf{have}(\mathbf{girl}, \mathbf{brush})$ and a propositional alternative to $\mathbf{have}(\mathbf{girl}, \mathbf{brush})$.

We assume, following von Fintel, that association with focus effects with *only* arise when resource domain selection for *only* and the presuppositions contributed by focus interpretation resolve to the same discourse topic. Under this account, subject-focus marking in (5b) introduces the presupposition that there is some accessible set of alternatives in the discourse model which is a subset of the focus semantic value $[[(5b)]]^f = \{\mathbf{have}(x, \mathbf{brush}) \mid x \in E\}$ including the ordinary semantic value $[[(5b)]]^o$ and an alternative proposition (Rooth 1992). (5b) is understood as discourse congruent with (5a) because $[[Q]]^o$, the ordinary semantic value of the discourse topical question *Which child has a brush?*, satisfies this presupposition, as indicated in (5f). $[[Q]]^o$ also provides a felicitous antecedent for the resource domain variable C for *only*, thus giving rise to the expected association-with-focus effect. We will argue that

(intonational) phrase. Focus assignment to a non-default nuclear stress position, as shown in (7), creates a conflict between the NSR and the FR. In this case, the FR rather than the NSR dictates the location of greatest prosodic prominence in the sentence (Selkirk 1995).

We propose that children who adopt the **AS** strategy resolve this conflict differently than adults by assigning greatest prosodic prominence to a syllable on the basis of the default stress pattern established by the NSR. As a result, these children assign greatest prosodic prominence to *kwast* ('brush') in both (3a) and (4a) and associate *only* with the object in both sentences on this basis. We show how this analysis accounts for our findings as well as the results of several other recent experiments with children.

References

- Paterson, Kevin B., Simon P. Liversedge, Caroline Rowland, & Ruth Filik (2003) Children's comprehension of sentences with focus particles. *Cognition* 89, 263-294
- Crain, Stephen, Weija Ni, & Laura Conway (1994). Learning, parsing and modularity. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum.
- Crain, Stephen, William Philip, Kenneth F. Drozd, Tom Roeper, & Kazuko Matsuoka (1992). Only in child language. Unpublished manuscript, University of Connecticut, Storrs, CT.
- Drozd, Kenneth .F. & Erik van Loosbroek (2006) The effect of context on children's interpretations of universally quantified sentences In V. van Geenhoven (Ed.) *Semantics in Acquisition*. Kluwer: Dordrecht von Fintel, Kai (1994) Restrictions on quantifier domains. UMASS Dissertation, Amherst, MA.
- Rooth, Mats (1992) A theory of focus interpretation. *Natural Language Semantics* 1, 75-116.

-
- Selkirk, Lisa (1995) Sentence prosody: Intonation, stress, and phrasing. In Goldsmith, John A. (Ed.) *The Handbook of Phonological Theory*. London: Blackwell.
- Pierrehumbert, Janet B. & Julia Hirschberg (1990) The meaning of intonation contours in the interpretation of discourse. In P.R. Cohen, J. Morgan, & M.E. Pollack (Eds.) *Intentions in Communication*. Cambridge: MIT Press. 271-311.

Crosslinguistic Studies on the Role of Information Structure for Ambiguity Resolution

Barbara Hemforth

Université de Provence

barbara.hemforth@lpl.univ-aix.fr

The information structure of a sentence finds itself reflected in the order of constituents as well as the prosodic contour of a sentence, among other things. In this paper, however, we will ask the question of in how far the information structural value of a constituent, given by such factors as ordering and prosody, influences preferences in the resolution of modifier attachment ambiguities, and how it interacts with anaphor resolution. In particular, we want to know in how far and in which way differences in the syntactic structure of different languages result in sometimes subtle, but sometimes drastic changes of interpretational preferences.

Let's first look at anaphor resolution: It has been shown here that the interpretation of pronouns very strongly depends on information structure with e.g., pronouns in German, English, and Finnish very often showing a preference for sentence topics and/or subjects, the preference being strongest when these factors go together. However, this preference does not show up across all languages, and given the specific ways of realizing anaphors in different languages, it should not do so. Across languages, we predict that the least explicit anaphor possible for a given construction should take the most salient antecedent. Using an anaphoric construction that is more explicit or lower on the accessibility hierarchy (Ariel, 1990), should be a cue to prefer a non-default antecedent.

If we compare German and French for constructions like (1, 2), which look very much alike across those two languages, we have to take into account different ways of realizing anaphoric expressions in those two languages.

- (1) Der Briefträger beobachtet den Straßenfeger, bevor er nach Hause ging.
- (2) Le facteur regarde le balayeur, avant qu'il rentre à la maison.
The postman is watching the street sweeper before he goes home.

For French, there exists an alternative to (2) involving a zero anaphor which can however only relate to the subject, which would be the default antecedent for the least explicit anaphor (3). Nothing comparable exists for German so that the least explicit anaphor in this language would be the pronoun.

- (3) Le facteur regarde le balayeur avant de rentrer à la maison.

In off-line questionnaires as well as in visual world experiments, we found a strong preference for the most salient antecedent (the subject) for German and a very strong preference for the object in French. Studies on English, where a zero anaphor alternative is possible but very rare in this particular construction according to corpus counts, show a German-like but much less pronounced subject preference (Hemforth, Scheepers, Colonna, Pynte, Konieczny, 2004).

To explain anaphor resolution across languages, we obviously have to take into account how information structure is realized in each particular language as well as where the particular anaphoric expression lies on the accessibility hierarchy for the particular construction and the particular language.

If now information structure and its realization in different languages determine ambiguity resolution, it would be interesting to see whether they have an effect on other non-anaphoric ambiguities as well. One type of ambiguity that

has attracted considerable crosslinguistic attention is the RC-attachment ambiguity. In sentences like (4), the relative clause “who died recently” can either modify the head of the NP (“son”; N1-attachment) or the more recent “teacher” (N2-attachment).

(4) Peter knew the son of the teacher who died recently.

It has been shown that many factors from all levels of linguistic processing exert an influence on the interpretation of this ambiguity, so it is a good testing ground for subtle effects. Hemforth, Fernandez, Clifton, Frazier, Konieczny, and Walter (2004) investigated RC-attachment preferences for German, English, and Spanish, varying the length of the relative clause (short vs. long) as well as the position of the NP-complex (preverbal subject or post verbal object). Across languages, we found an effect of length, which is usually attributed to prosodic factors (Fodor, 1998; but see below), but we also established crosslinguistic differences with respect to position effects: In German and Spanish, the number of N1-attachments was strongly reduced for NP-complexes realized as preverbal subjects compared to post verbal objects. However, this was not the case for English. We argued that the fact that the preverbal position in German and Spanish does not obligatorily have to be filled with the subject makes this position more informative with respect to “givenness” than in English where there is actually not much of a choice. Assuming furthermore, that in a null context the supposedly given subject does not require or not even attract much information to accommodate the discourse structure, a less pronounced attraction of relative clauses can be explained. This hypothesis predicts that in languages with rather fixed word order similar to English, we should find no positional effects, whereas in languages allowing for more variation in particular

with respect to default topic positions, position dependent interpretational preferences should show up.

We tested these predictions for French (fairly fixed constituent ordering) and Hebrew (fairly free constituent ordering). As predicted, for close translations of the materials used for German, English, and Spanish, we did not find a position effect for French (Hemforth et al., 2006), whereas we found a pattern very close to German and Spanish for Hebrew (Shuval & Hemforth, 2006).

A pattern, however, that has been established for all languages investigated so far, including French and Hebrew, is the length effect: longer relative clauses tend to be attached to N1 more often than short ones. It has been argued (Fodor, 1998) that this is due to prosodic balancing which is assumed to play its role in silent reading as well as in spoken language (Implicit Prosody Hypothesis). However, the preference to attach a long modifier to a more central element of a sentence may have its basis in its enhanced informational load as well (Almor, 1999). Assuming that more information means more processing effort, this enhanced effort has to be justified, and it is more justified for the more central elements, i.e. direct arguments of a proposition (Frazier & Clifton, 1996). In a series of experiments on French and German we tried to disentangle a prosody-based hypothesis from one based on information load (Hemforth, Petrone, D'Imperio, Pynte, & Colonna, 2006). We looked at PP-attachment ambiguities like (5, 6) where we used short and long city names.

- (5) L'assistant de l'avocat de a. Apt / b. Albertville / c. Aix-en-Provence a discuté avec le nouveau juge.
The assistant of the lawyer from discussed with the new judge.
- (6) Le nouveau juge a discuté avec l'assistant de l'avocat de a. Apt / b. Albertville / c. Aix-en-Provence.
The new judge discussed with the assistant of the lawyer from

Changing only the number of syllables in the PP without adding content words did not have any effect on attachment preferences. Adding content by inserting adjectives as in (7) however increased the number of N1-attachments reliably.

- (7) Le nouveau juge a discuté avec l'assistant de l'avocat de a. Apt / b. la belle ville d'Apt.
The new judge discussed with the assistant of the lawyer from Apt / the beautiful city of Apt.

In an experiment on German, we directly compared constructions of identical syllable length, just lengthening the city name in one case and adding an adjective in the other one as in (8).

- (8) Der neue Richter diskutierte mit dem Assistenten des Notars aus a. Au / b. Au-Winterach / c. dem schönen Au.
The new judge discussed with the assistant of the lawyer from Au / Au-Winterach / the beautiful Au.

Only lengthening by a content word increased the number of N1-attachments, strongly suggesting that it is actually the amount of information to be processed which is at the basis of the length effect in modifier attachment.

To sum up these experiments, we can conclude that we can predict crosslinguistic ambiguity resolution when taking into account the explicitness of information given (e.g., for anaphoric expression), the amount of information presented (information load), and the way information structure is realized in a particular language.

References

- Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, 106, 748-765.
- Ariel, Mira (1990) *Accessing noun-phrase antecedents*. London: Routledge.
- Fodor; J.D. (1998). Learning to parse? *J. of Psycholing. Res.*, 27, 285-319.
- Frazier, L., & Clifton, C. (1996). *Consrual Theory*. Boston: MIT Press.
- Hemforth, B., Fernandez, S., Clifton, C., Frazier, L., Konieczny, & Walter, M. (2006). Manuscript submitted.
- Hemforth, B., Petrone, C., D'Imperio, M., Pynte, J., & Colonna, S. (2006). Poster presented at CUNY, March 2006, New York.
- Hemforth, B., Scheepers, C, Colonna, S., Pynte, J., & Konieczny, L. (2004). Poster presented at AMLaP, Sptember 2004, Aix en Provence.
- Shuval, N., & Hemforth, B. (2006). Submitted to AMLaP, August 2006, Nijmegen.

Information Structure and OV Order

Thorbjörg Hróarsdóttir

University of Tromsø

Thorbjorg.Hroarsdottir@hum.uit.no

The aim of this paper is to argue that the cue for the OV/VO parameter was expressed through information structure in Older Icelandic (OI). I will show how a change in encoding of information structure may lead to change in the basic order of verb and its complement.

Roberts (1997) argues that English lost overt object-movement due to the loss of morphological case in Middle English. As seen from the English viewpoint, low-level facts of inflectional morphology may express the relevant cue for parameters and so the loss of inflection may (but does not have to) lead to a grammar change. I will show how this analysis does not carry over to Icelandic as the loss of OV in Icelandic took place despite rich case morphology. I assume that learners must watch out for a universal cue for the positive value of the OV/VO parameter. However, this cue may be expressed differently among the languages: while it may have been expressed through morphology in Old English (OE), it was expressed through information structure in OI. In both cases, external effects led to fewer expressions of the relevant cue and a grammar change took place.

This predicts that the relationship between strong morphology and OV order has been overestimated. I claim that this relation is only one-directional, in the way that languages that lose strong morphology also tend to lose overt OV order (English), because of structure sandwiching, but not vice versa. Therefore, languages may have overt OV orders despite weak morphology, and languages may lose overt OV structure regardless of whether there have been any changes

in the morphology or not (Icelandic). Moreover, many modern languages have rigid word order despite their case morphology, and vice versa (cf. Siewierska 1998).

The OV/VO word order patterns in OI are stable until the seventeenth century, where there is a gradual loss of the OV word order patterns for two centuries, or, in other words, a gradual increase in language use to have focused elements in postverbal (VO) position (a shift in discourse property). During the seventeenth and eighteenth centuries, a gradual change took place in the history of Icelandic, paving the way for the parameter change in the beginning of the nineteenth century, when we have a change from a grammar allowing the variation of both OV and VO word order patterns, to a grammar allowing only (pure) VO word order. Hence, DPs expressing new information were more often put in postverbal position, with the consequences that there was a gradual drop in the frequency of OV word order patterns until OV had dropped below a certain threshold to be useful as cues (Lightfoot 1999). In other words, this increased frequency of focused DPs in postverbal position gradually led to fewer expressions of the relevant cue. Hence, there is a(n abrupt) parameter change within the next generation acquiring the language (in the beginning of the nineteenth century).

In sum, this means that a shift in discourse property, or simply a change in language use, can lead to a grammar change; a change in the basic word order in the I-language.

A central question in the comparison of OV- and VO-languages is whether the difference results from having more object movements in OV-languages, or more verb movements in VO-languages. I argue, following Hróarsdóttir (2000), that although a uniform VO-base hypothesis with overt versus covert leftward movement of objects yields positive results in the synchronic analysis of the

Older Icelandic VP, it cannot easily handle the diachronic aspect. The displacement of the verb could be a consequence of either (i) head movement of the verb (Haider 1992), or (ii) movement of a larger constituent containing the verb. e.g. VP (Hinterhölzl 1997; Kayne 1998; and Hróarsdóttir 2000). I will assume the latter option here.

Finally, I will show how Object Shift in Modern Icelandic supports the claim that the cue for OV in OI was expressed through information structure. In order to be able to shift, the DP object must represent old information, while a DP object representing new information cannot shift and must remain internal to the VP. Object Shift can now be regarded as a movement of the object along with the VP, in front of the negative adverbial. While full DPs must move out of the VP (to get their case checked in [Spec, AgrOP] in the functional domain), unstressed DPs and pronouns do not have to exit the VP (because they do not have strong enough features). This connects to the claim that objects expressing new information in OI must move out of the VP, prior to a remnant VP-movement, and hence end up postverbally, while light objects do not have to exit the VP, and therefore they may move along within the preposed VP (containing the finite verb) and thereby end up preverbally.

References

- Haider, Hubert.(1992). “Branching and discharge”. *Arbeitspapiere des Sonderforschungs-bereichs 340* 23. University Stuttgart. [Revised version in *Lexical Structure, Proceedings of the 1991 Utrecht Conference*, P. Coopmans & J. Grimshaw (eds.)]
- Hinterhölzl, Roland.(1997). A XP-movement account of restructuring. Ms., University of Southern California.
- Hróarsdóttir, Thorbjörg. (2000). *Word Order Change in Icelandic. From OV to VO*. [Linguistik Aktuell/Linguistics Today 35]. Amsterdam: John Benjamins.

-
- Lightfoot, David W.(1999).*The Development of Language: Acquisition, Change, and Evolution*. Oxford and Malden, Mass: Blackwell Publishers.
- Roberts, Ian G.(1997).“Directionality and word order change in the history of English”. In Ans van Kemenade and Nigel Vincent (eds.):*Parameters of Morphosyntactic Change*, 396-426.Cambridge:Cambridge University Press.
- Siewierska, Anna.(1998).“Variation in major constituent order: A global and a European perspective”.In Anna Siewierska (ed.): *Constituent Order in the Languages of Europe*, 475-551.Berlin: Mouton de Gruyter.

Is There a Focus Position in Aghem?

Larry Hyman

Maria Polinsky

UC Berkeley

UC San Diego

hyman@berkeley.edu

polinsky@ling.ucsd.edu

This paper presents and analyzes the expression of focus in Aghem, a Western Grassfields Bantu language of the Ring Subgroup, known for its particular richness in focus marking (Anderson 1979, Hyman 1979a, Watters 1979). Aghem noun phrases appear to have different marking which often correlates with the information- structural status of the respective NP. Hyman (1979a, 1985, 2006) presents a detailed analysis of the internal syntax of the DP in Aghem, and we take those results as a given. In this talk, we will concentrate on the interaction between independently motivated clause types (reflected in the form of DPs in particular; cf. the object DP in (1)-(3)) and information structure. More specifically, we address the motivation for a dedicated focus projection in Aghem.

Aghem (and Bantu languages in general) seems to show a strong correlation between focus/WH interpretation and a particular syntactic position — immediately after the verb (IAV), cf. (4)-(7). Researchers have proposed that this position is uniquely identified with focus (Horvath 1986), which has in turn been taken as the motivation for expanded left periphery (Aboh 2006), shown in (8).

We present empirical arguments against the cartographic representation of focus (8), and then show that it is impossible to maintain a direct link between the IAV and focus encoding. The apparently focused interpretation of the IAV follows from the range of interpretations available to the constituent occurring in that position and thus depends on the category of that constituent (focus for

indefinite expressions, including *wh*-words which belong to that class in Aghem; ambiguity between focus and non-focus for postverbal DPs, and non-focus for pronouns). In addition to this result, the IAV, while linearly definable on the surface on the surface, is not the same position structurally — while arguments must appear as sister to the verb, PPs and adverbials can right-adjoin to the lower *vP* for focus interpretation. We show that focus interpretation is determined scopally over a constituent inside the *vP*, bound by the overt focus operator or interrogative operator in *spec,CP* (9). This binding relation, motivated for many languages in similar constructions, is quite standard and does not require movement operations or a dedicated focus projection.

Examples:

- (1) bvú $\downarrow\text{tí}$ mò zí kí-bé $\text{á}\downarrow\text{zóó}$
 dogs D PAST₂ eat D-fufu yesterday
 ‘The dogs ate fufu yesterday.’
- (2) bvú $\downarrow\text{tí}$ mò bé $\downarrow\text{kí}$ zí $\text{á}\downarrow\text{zóó}$
 dogs D PAST₂ fufu D eat yesterday
 ‘The dogs ate fufu YESTERDAY.’
- (3) bvú $\downarrow\text{tí}$ mò zí $\text{á}\downarrow\text{zóó}$ $\text{bé-}\downarrow\text{kó}$
 dogs D PAST₂ eat yesterday fufu-OBL
 ‘The dogs ate fufu YESTERDAY.’
- (4) bvú $\downarrow\text{tí}$ mò ńńη nô
 dogs D PAST₂ run FOC
 ‘The dogs ran.’
- (5) a. à mò ńńη tɪbvú
 EXPL PAST₂ run D-dogs
 ‘The DOGS ran.’

- b. à m̀ ñɨŋ ńdú↓ghó (à)
EXPL PAST₂ run who INTERR
'Who ran?'
- (6) a. bvú ↓tɨ m̀ zí kí-bé (nò)
dogs D PAST₂ eat D-fufu FOC
'The dogs ate FUFU.'
- b. bvú ↓tɨ m̀ zí kẁ (à)
dogs D PAST₂ eat what INTERR
'What did the dogs eat?'
- (7) a. bvú ↓tɨ m̀ zí á↓zóó bé-↓kó
dogs D PAST₂ eat yesterday fufu-OBL
'The dogs ate fufu YESTERDAY.'
- b. bvú ↓tɨ m̀ zí zɨn bé-↓kɔ
dogs D PAST₂ eat when fufu-OBL
'When did the dogs eat fufu?'
- (8) ForceP > TopicP > FocP (> TopicP) > FinP ... > TopicP > FocP > VP
(cf. Rizzi 1997, 2002, Beletti 2002)
- (9) [_{CP} Op_i [_{C'} [[_{TP} ... XP_i ...] nò/à]]]
[+FOC]/[+INTERR]

References

- Aboh, Enoch. 2006. If we see Focus, you go left and I go right! *Paper presented at the International Conference on Bantu Grammar*. SOAS, London, April 2006.
- Anderson, Stephen C. 1979. Verb structure. In Hyman (1979a), 73-136.

-
- Belletti, Adriana. 2002. Aspects of the low IP area. In Luigi Rizzi (ed.). *The structure of IP and CP: The cartography of syntactic structures 2*. Oxford: Oxford University Press.
- Hyman, Larry M. (ed.). 1979a. *Aghem grammatical structure*. Southern California Occasional Papers in Linguistics 7. Los Angeles: University of Southern California.
- Hyman, Larry, ed. 1979b. *Aghem grammatical structure*. (SCOPIIL 7.) Los Angeles: USC Dept of Linguistics Publications.
- Hyman, Larry. 1985. Dependency relations in Aghem syntax: The mysterious case of the empty determiner in Aghem. *Studies in African Linguistics, Supplement 9*: 151-156.
- Hyman, Larry. 2006. Focus marking in Aghem: Syntax or semantics? Paper presented at the Conference on Focus in African languages, Berlin, October 2005 and submitted to the proceedings.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.). *Elements of grammar*. Dordrecht: Kluwer.
- Rizzi, Luigi. 2002. Locality and left periphery. In Adriana Belletti, (ed.). *Structures and beyond. The cartography of syntactic structures*, vol. 3. Oxford: Oxford University Press.
- Watters, John R. 1979. Focus in Aghem: a study of its formal correlates and typology. In Hyman (1979b), 137-197

Why Negative Adverbials of Degree, Manner, and Frequency Must be Focussed

Katalin E. Kiss

Linguistic Institute of the Hungarian Academy of Sciences

ekiss@nytud.hu

1 The problem

The talk proposes an answer to an old question of Hungarian syntax: why positive adverbials of degree, manner, and frequency, and their negative counterparts have different word order possibilities. Whereas positive adverbials must (or can) be adjoined to PredP, their negative counterparts must be focussed:

- (1)a. János [_{PredP} **nagyon** [_{PredP} el-fáradt]] (2)a. *János [_{PredP} **alig** [_{PredP} el-fáradt]]
John very prt-got.tired John hardly prt-got.tired
b. *János [_{FocP} **NAGYON** fáradt [_{PredP} el]] b. János [_{FocP} **ALIG** fáradt [_{PredP} el]]
- (3)a. János [_{PredP} **gyakran** [_{PredP} el-késik]] (4)a. *János [_{PredP} **ritkán** [_{PredP} el-késik]]
John often prt-is.late John seldom prt-is.late
b. János [_{FocP} **GYAKRAN** késik [_{PredP} el]] b. János [_{FocP} **RITKÁN** késik [_{PredP} el]]

One and the same adverbial may require different positions depending on whether it is used as a positive adverbial of degree, or a negative adverbial of manner:

- (5)a. Jánost [_{PredP} **csúnyán** [_{PredP} meg-verték]]
John-acc badly prt-beate-they
'John was badly beaten.'
b. János [_{FocP} **CSÚNYÁN** vasalta [_{PredP} ki az inget]]
John badly ironed prt the shirt
'John ironed the shirt badly.'

NPs modified by positive and negative indefinite numerals show the same contrast:

- (6)a. János [_{PredP} sok példát [_{PredP} meg-oldott]]
 John many examples prt-solved
 b. János [_{FocP} SOK PÉLDÁT oldott [_{PredP} meg]]

- (7)a. *János [_{PredP} kevés példát [_{PredP} meg-oldott]]
 John few examples prt-solved
 b. János [_{FocP} KEVÉS PÉLDÁT oldott [_{PredP} meg]]

2 A key to the mystery: numerically modified NPs

The explanation to be proposed is based on the analysis of the behavior of noun phrases modified by definite numerals. They can stand either adjoined/internal to PredP (8a,b) or in focus position (9), and the two positions are associated with clearly different interpretations:

- (8)a. János [_{PredP} két doboz cigarettát [_{PredP} el-szívott naponta]]
 John two packet cigarettes prt smoked daily
 'John smoked (*at least*) 2 packets of cigarettes a day.'
 b. János [_{PredP} el-szívott naponta két doboz cigarettát]
 John prt smoked daily two packet cigarettes
 'John smoked (*at least*) 2 packets of cigarettes a day.'

- (9) János [_{FocP} KÉT DOBOZ CIGARETTÁT szívott [_{PredP} el naponta]]
 'John smoked (*exactly*) 2 packets of cigarettes a day.'

The meaning difference is particularly conspicuous under negation:

(10)a. János [_{NegP} nem szívott [_{PredP} el naponta két doboz cigarettát]
 John not smoked prt daily two packet cigarettes
 'John didn't smoke 2 packets of cigarettes a day (*he smoked less*).'

b. János [_{NegP} nem [_{FocP} KÉT DOBOZ CIGARETTÁT szívott [_{PredP} el naponta]]
 'It wasn't 2 packets of cigarettes that John smoked a day (*it was more or less*).'

The basic meaning of a numeral modifier n in natural language is 'at least n ' (even if a so-called scalar implicature often results in the interpretation 'at least n and at most n '). This well-known semantic property of numeral modifiers holds for scalar modifiers, in general. Thus any scalar element can be interpreted as upward entailing; in other words, the truth of a sentence with a scalar modifier remains unchanged also in a situation in which the value of the given scalar element is replaced by a higher value. In focus position, however, the upward entailment of a scalar element is blocked (cf. also Fretheim (1992) and van Kuppevelt (1996)). The fact that in focus position, a scalar modifier loses its capability of upward entailment will be derived from the function of focus, expressing exhaustive identification, i.e., excluding possible alternatives.

It will be shown that in the case of scalar elements in the negative domain of a bidirectional scale, e.g. in the case of *few*, the upward entailment associated with scalar modifiers leads to a semantic anomaly. Compare:

- (11)a. I have read **many** 18th century English novels.
 b. I have read **few** 18th century English novels.

(11a) remains true also if I have read all 18th century English novels. However, in the case of (11b), the replacement of the value of *few* with that of *many* or *all* would replace a negative scalar value by a positive one, contrary to the speaker's intentions.

The upward entailing interpretation of scalar elements in the negative domain of bidirectional scales is prevented by their obligatory focussing – cf.

- (12)a. [_{FocP} **Kevés 18. századi angol regényt** olvastam [_{PredP} el]]
 few 18th century English novel-acc read-I prt
 'I have read few 18th century English novels.'
 b.* [_{PredP} **Kevés 18. századi angol regényt** [_{PredP} el olvastam]]

3 Scalar adverbials

It will be claimed that adverbials of degree, manner, and frequency are also scalar elements. In the case of those located in the positive domain of a bidirectional scale, upward entailment has no unwelcome consequences – see e.g. (13a). In the case of negative adverbials, on the other hand, the upward entailing interpretation must be excluded by the focussing of the adverbial – see (13b).

- (13)a. János [_{PredP} **nagyon** [_{PredP} meg-sérült]]
 John very prt was.hurt
 'John was hurt very much.'
 b. János [_{FocP} **ALIG** [_{PredP} sérült meg]]
 John barely was.hurt prt
 'John was barely hurt.'

Positive adverbials of manner and frequency can also be focussed. The meaning difference between a non-focussed and a focussed positive adverbial consists in the presence vs. absence of upward entailment. (15b) is incoherent because focussing blocks the upward entailment required by the context:

- (14)a. János **elég jól** ki-töltötte a tesztet. 100-ból 100 pontot ért el.
 John quite well out-filled the test Of 100 points, he achieved 100 points.
 b.??János **ELÉG JÓL** töltötte ki a tesztet. 100-ból 100 pontot ért el.
 John quite well filled out the test Of 100 points, he achieved 100 points.

References

- Fretheim, T. (1992) The effect of intonation on a type of scalar implicature. *Journal of Pragmatics* 24, 1-30.
- Kuppevelt, J. van (1996) Inferring from topics. *Linguistics and Philosophy* 19, 393-443

Scalar Implicatures – Which Theory Best Explains Their Processing and Acquisition?

Tanya Reinhart

Tel Aviv University

reinhart@post.tau.ac.il

Scalar implicatures have become a central topic of research since 2000 and there are many competing analyses of their semantics and computation. Taking Chierchia's (2004) perspective, a central debate is between the "pragmatic", or context-driven approaches, and the computational system (CS) approaches that place their computation in semantics (Chierchia) or syntax (Fox). On the first view, the identification and interpretation of implicatures is context-dependent. Hence, the computation must be global, enabled only when the contextual trigger is encountered. However, Landman 2000 and Chierchia 2004 showed that global computation is infeasible semantically. In Chierchia's analysis, the computation of implicatures is strictly compositional and local, determined by CS entailments definitions. To overcome the problem of the contextual identification of implicatures, Chierchia assumes that scalar implicatures are associated with scalar items by default, starting at the lexicon, but they can be cancelled by context.

In many cases, it appears that such debates may be more conceptual than empirical. Furthermore, within the approaches locating implicatures in the CS, some of the more intricate competing analyses are not easily distinguishable from one another, on empirical grounds. I will focus here on one empirical criterion for theory selection. I argue in *Interface Strategies* that within Chomsky's hypothesis of optimal design, the computations postulated by CS theory may have direct correlates in the processing of derivations.

Correspondingly, facts about processing may be relevant for deciding between competing theories. Specifically, I argue there that the relatively rare finding of 50% group performance in acquisition is an indication of a processing difficulty. It is found when the required computation exceeds working memory capacity. Since children's working memory is less developed than adults, tasks that are within the processing ability of adults may exceed children's ability. Performance at the range of 50% has been found so far in areas that require global optimality computations – the construction and comparison of competing derivations (as in coreference and shifted focus computations). The stronger hypothesis is that whenever this range of performance is found, some global optimality is at work, or at least, the source of the processing difficulty should be identifiable in the properties of the computation required by the CS.

As of the early 2000s, experiments repeatedly found the 50% range of performance in the acquisition of scalar implicatures. Chierchia et al. 2001 and Gualmini et al. 2001 showed that in downward entailing (DE) contexts, as (1), where no implicature is involved, children perform essentially like the adult control group, judging the sentence as true in a situation where some of the dwarfs chose both a banana and a strawberry.

- (1) Every dwarf who chose a banana or a strawberry received a jewel.
- (2) Every boy chose a skate-board or a bike.

However, in the implicature case (2), uttered in the same story context, the adult control group rejected it in 100% of the trials, while children performed at the 50% range (accepting the target sentence about half of the times). Essentially the same range of performance was found in the implicatures context for the *{some...all}* and the *{start, finish}* scales in Papafragou and Musolino 2002,

2003. These findings do not extend to numeral scalar items, where children perform much closer to adults. The question then, is which of the competing approaches can explain these acquisition findings.

Under Chierchia's analysis, implicatures are inserted into the semantic derivation by default, regardless of whether they are eventually realized or not. For any English expression α , $\|\alpha\|$ is its standard semantic value, which Chierchia calls the "plain" value. But α is also assigned compositionally a scalar value - $\|\alpha\|_S$, which is computed based on the scale of one of the elements in α . The plain value of (2) corresponds to *Every boy* ($\lambda x(x \text{ chose a banana or a strawberry})$), and its scalar value is (3).

- (3) Every boy ($\lambda x(x \text{ chose a banana or a strawberry} \ \& \ \neg x \text{ chose a banana and a strawberry})$)
- (4) Every dwarf ($\lambda x((x \text{ chose a banana or a strawberry}) (x \text{ received a jewel}))$)

When the disjunction occurs in the (restrictive) scope of a DE operator, the implicature clause is removed, a procedure governed by "the strength condition", informally summarized in (5).

- (5) Strength condition: The scalar value of α cannot be weaker than its plain value (where a representation α is stronger than a representation β iff α entails β .)

(5) is checked locally, at each step of the derivation, and its effect is to filter out implicatures when they lead to weakening of the original information content. The semantic derivation of (1) starts precisely the same way as (2), i.e. a predicate similar to (3) is initially constructed. However, when the DE operator of (1) is encountered, the plain value (4) entails the scalar value, which

would be *Every dwarf* ($\lambda x((x \text{ chose a banana or a strawberry} \ \& \ \neg x \text{ chose a banana and a strawberry}) \ (x \text{ received a jewel}))$). Hence, the scalar value is weaker than the plain value, and it is filtered out.

However, this elegant analysis fails to explain the acquisition findings. First, it is not clear why children should have problems with processing active implicatures, as (2). The proposed computation is strictly compositional and involves just local optimality evaluation of alternatives. Based on other areas of acquisition, there is no reason to suspect that children have massive problems with this type of computation. Suppose, nevertheless, that children have problems even with the local application of the strength condition. The crucial next question is why the same does not hinder them in (1). Filtering out the implicature here requires precisely the same local optimality procedure of comparing the plain and the scalar value. While in (2), which children fail to process, this computation applies only once, in (1), it applies twice, so it should have been harder, or at least equally hard for children to process.

The source of the problem is Chierchia's assumption that implicatures are associated with scalar items by default (following, in this respect, a long neo Gricean tradition starting with Gazdar 1979 and Horn 1984). The alternative context-driven view is that the default interpretation is the plain value, and scalar implicatures, like the Gricean particularized conversational implicatures, are generated only if there is contextual reason to do so. (E.g. Breheny, Katsos and Williams in press.) If so, then in contexts like (1), children would not even attempt to compute an implicature. The problems they have with computing implicatures like (2) suggest that some global optimality is involved when implicature assessment has to apply. Nevertheless, Chierchia has proved that the semantic computation of implicatures cannot be global in the sense of the context-driven approaches.

I argue that the actual semantic computation of implicatures is as defined in (any of the variants) of the CS approach. However, as in the context-driven approach, in the standard cases, it is triggered only by the context. Hence, when required by the context, a derivation that has been completed under its plain value has to be reopened and an alternative with the scalar value must be constructed compositionally. The two representations must then be compared against the context. This has all the properties of a costly global optimality computation (reference-set), that children are known to be unable to perform.

Given that the scalar computation itself is available in UG, it is in principle possible that there are areas where it applies indeed during the derivation. Suppose that with numerals, Chierchia's procedure is lexicalized, i.e. it does apply as default. A numeral, then, is always computed locally against its scalar alternatives, carrying the 'at most' implicature through the derivation, unless it is cancelled by a DE operator, or by other contextual factors. As we saw, this type of computation should not cause a processing crash with children. So this would explain why they have no problem with numeral scalar implicatures.

References

- Breheny, Richard, Napoleon Katsos and John Williams. in press. Are Generalised Scalar Implicatures Generated by Default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*.
- Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (ed.), *Structures and Beyond*. Oxford University Press, Oxford.
- Chierchia, G, S. Crain, M.T. Guasti, A. Gualmini and L. Meroni. 2001. The acquisition of disjunction: evidence for a grammatical view of scalar implicatures. In *Proceedings of the 25th Annual Boston University Conference on Language Development*, ed. Anna H.-J. Do, et al., 157-168. Somerville, Mass.: Cascadilla Press.

-
- Gazdar, G. A. (1979). *Pragmatics: Implicatures, Presupposition and Logical Form*. New York: Academic Press.
- Gualmini, A., S. Crain, L. Meroni, G. Chierchia and M.T. Guasti. 2001. At the semantics/pragmatics interface in child language. In *Proceedings from Semantics and Linguistic Theory 11*, 231-247. Ithaca, NY: Cornell University.
- Horn, Larry. 1984. Toward a new taxonomy for pragmatic inference. In D. Schiffrin (ed.) *Meaning, Form and Use in Context: Linguistic Applications (GURT '84)* (pp. 11-42). Washington D.C.: Georgetown University Press.
- Landman, Fred. 2000. *Events and Plurality*. Dordrecht: Kluwer
- Papafragou, A., and J. Musolino. 2002. The pragmatics of number. *Proceedings from the 24th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, 2002..
- Papafragou A. and J. Musolino (2003) Scalar Implicatures: Experiments at the Semantic-Pragmatics Interface. *Cognition*, 80, 253-282

S-focus and Relativized Stress F

Mats Rooth

Cornell University

mr249@cornell.edu

This talk will analyze the hypothesis that the phonological realization of focus marking is subject to some version of a Stress-F constraint, and how this constraint interacts with the grammatical representation of second-occurrence focus (S-Focus).

Stress F

Let B be an F-marked phrase with scope P. Then the strongest stress in the phonological realization of P falls within the realization of B.

A couple of motivations for this constraint have been given. Truckenbrodt (1995) argued that in 'farmer examples' like (i), the phonological domain in which the focus is maximally prominent matches the semantic scope of the focus.

- (i) An American farmer told a Canadian farmer a joke.

This motivates the reference to P and its phonological realization in the statement of Stress F.

Selkirk (2002 and elsewhere) has suggested that there are prominence differences between alternative-semantics F which is subject to some version of Stress F, and new-information focus. In this account, [Bobby] in (ii) has an alternative-semantics focus F which is subject to Selkirk's version of Stress F, while [Ashley] has a new-information.

-
- (ii) You know what? You only introduced Ashley_f to Bobby_F yesterday.
(You didn't introduce Ashley to John yesterday.)

While both F and f in this example are accented, F is claimed to be more prominent, and this is attributed to Stress F.

Finally, Rooth (1996) suggested that Stress F is responsible for the special realization of second-occurrence focus.

- (iii) Even Eva_F only gave copies to the graduate_SOF students.

The hypothesis is that it is characteristic of an SOF configuration that the F has wider semantic scope than the f. In interaction with the the Stress F constraint, this results in greater stress on the realization of F than the realization of SOF, which is supposed to be responsible for the phonological and phonetic realization of SOF.

I'll look at two variant formulations of Stress F. Overt-antecedent Stress-F is motivated by the claim in Rooth (1996) that SOF phonology only shows up when the broader-scope F has an overt antecedent (is anaphoric). Relativized stress F is motivated by examples like (iv), where one can argue that two applications of the simple Stress F constraint impose contradictory constraints, because the phonological domains are nested.

- (v) You know what? You only introduced Mona to Bobby_F yesterday. You also only introduced Ashley_F to Bobby_SOF yesterday.
- (vi) Relativized Stress-F let B be an F-marked phrase with scope P Then the strongest stress in B is strictly stronger than any stress in P which is not contained in an F-marked subconstituent of P whose scope is at least P.

The problem that (v) poses for simple Stress F is curiously reminiscent of the first operator effects in alternative semantics focus, namely that a focus is

<http://www.sfb632.uni-potsdam.de>

predicted to interact with the first focus-sensitive operator up the tree. I will sketch ideas for making something of this by closely tying prosody to the recursive semantics of focus. I will also review the discussion of the semantic first-operator effect in Rooth (1996a) and Krifka (200x).

Finally, I will look at data which challenge the validity of the hypothesis that the semantic scope of F's matches relative phonological prominence. In the rice example (vii), it turns out that this hypothesis predicts the wrong locus for SOF phonology, i.e. on [eat] rather than [rice]. In fact it can be shown that [rice] in this configuration has SOF properties.

(vii) People who grow rice generally only eat rice.

Contrastive Topics

Satoshi Tomioka

University of Delaware

stomioka@udel.edu

The notion of contrastiveness is connected to diverse linguistic phenomena (e.g., exhaustive answers in question-answer pairs, contrastive statements like *A but not B*, association with focus with *only*, *always* etc.). This presentation primarily concerns one sub-case of expressing contrastiveness; *Contrastive Topics*, which are often described as expressions of uncertainty, partiality, or limited knowledge. Most of the previous formal semantic/pragmatic analyses of contrastive topics come from the investigations of A- vs. B-accent in English (Jackendoff 1973, Roberts 1994, Buring 2003) and Topic accents in German (Krifka 1998, Buring 1997, 2003). I will not make full reference to these phenomena but concentrate instead on contrastive topics in Japanese, which share similar semantic and pragmatic properties with those relatively well-studied phenomena but display more liberal, perhaps unexpected, distributional patterns.

Japanese contrastive topics have two characteristics that are morphologically/ phonologically observable; (i) it is overtly marked with the particle *wa*, and (ii) high pitch falls onto either *wa* itself or the phrase *wa* attaches to, which is accompanied by the reduction of pitch accents of the material that follows the contrastive topic. This second pattern is reminiscent of a proto-typical focusing strategy (cf. Nagahara 1994, Ishihara 2002, 2003 among others). An ideal analysis would, therefore, derive the meaning of Japanese contrastive topics by combining the meaning of the morpheme *wa* with the contribution of high pitch accent placed on them.

What does the contrastive *wa* mean, then? In many ways, it evokes the same kind of partiality or unresolvedness that B-accent in English does. For instance, a partial answer to a Wh-question is expressed with a contrastive topic, as in (1a). The distribution of the contrastive *wa*, however, goes far beyond this well-documented contexts; it can appear in interrogative sentences themselves ((1b)) or in the imperative mood ((1c)).

- (1) a. (Who left?) KEN-wa kaet-ta
 Ken-CT leave-past
 ‘(At least) Ken left.’
- b. ... Zyaa KEN-wa doko-e it-ta-no
 then Ken-CT where-to go-past-Q
 ‘Where did KEN go?’
- c. SUUGAKU-wa benkyoo-site-ok-e
 math-CT study-do-prepare-IMP
 ‘At least, study MATH.’

Other non-assertion speech act sentences that can host contrastive topics include exhortitives and performatives.

My analysis is inspired by the one offered in Hara (2006), who provides many of the crucial ingredients for computing uncertainty/ partiality. The main components of the analysis I put forth are the following.

- High pitch accent on a contrastive topic elicits a set of alternatives. However, the members of this set can be ‘ordered’ in terms of strength (cf. Sauerland’s 2004 notion of ‘scaler alternatives’).
- Contrastive operator (CON in Hara 2006) is adjoined to a Speech Act constituent.

-
- The use of a contrastive topic presupposes that among the set of alternative speech acts (= the ‘focus value’ of the speech act constituent that CON is attached to), there be a speech act that is ‘stronger’ than the actual speech act (= the ‘ordinary’ value of the speech act constituent).
 - The Gricean reasoning provides that there be a good reason why the speech act agent didn’t perform the stronger speech act.

The advantages and disadvantages of this ‘speech-act-based’ analysis, particularly in comparison with some of the knowledge-based analyses (e.g., Hara 2006 and van Rooij and Schulz 2004), will be discussed. One of the major issues is that our semantics/pragmatics of CON does not specify the reason why a weaker speech act was performed: It could be that the agent’s knowledge was partial/incomplete, as often described in the knowledge-based analyses, but it could also be that the agent with complete knowledge simply wished the information to be kept from being public or she believed that performing the stronger speech act was impolite or socially unacceptable. It will be shown that the actual uses of contrastive *wa* are more in accordance with the speech-act-based analysis. Other curial issues to be discussed are; the possibility of *wa* in embedded contexts and its correlation with the embedding of speech acts (cf. Portner 2005, Krifka 2001, 2002), the relation between the thematic and the contrastive topics, which share the same morphology.

References

Buring, Daniel (1997) ‘The great scope inversion conspiracy,’ *Linguistics and Philosophy* 20: 175-194.

-
- Buering, Daniel (2003) ‘On D-trees, beans, and B-accents,’ *Linguistics and Philosophy* 26: 511-545.
- Hara, Yurie (2006) *Grammar of Knowledge Representation: Japanese Discourse Items at Interfaces*, Doctoral Dissertation, University of Delaware.
- Ishihara, Shin-ichiro (2003) *Intonation and Interface Conditions*, Doctoral Dissertation, MIT.
- Krifka, Manfred (1998) ‘Scope inversion under the rise-fall pattern in German,’ *Linguistic Inquiry* 29: 75-112.
- Krifka, Manfred (2001) ‘Quantifying into question acts,’ *Natural Language Semantics* 9: 1-40.
- Roberts, Craige (1996) ‘Information structure in discourse: Towards an integrated formal theory of pragmatics’, in J. H. Yoon and Andreas Kathol (eds.), *OSU Working Papers in Linguistics 49: Papers in Semantics*, 91–136.
- Sauerland, Uli (2004) ‘Scalar implicatures in complex sentences,’ *Linguistics and Philosophy* 27: 363-391.

A Lexical Perspective on Discourse Structure and Semantics

Bonnie Webber

University of Edinburgh

bonnie@inf.ed.ac.uk

D-LTAG is a *lexicalized* approach to discourse relations, providing an account of how lexical elements (including phrases) anchor discourse relations and how other parts of the text provide arguments for those relations.

D-LTAG arose from a belief that the mechanisms for conveying discourse relations were unlikely to be entirely different from those for conveying relations within the clause. Because the latter can be anchored on lexical items, D-LTAG was developed as a lexicalized grammar for discourse-specifically, a lexicalized Tree Adjoining Grammar (Schabes, 1990). A lexicalized TAG (LTAG) differs from a basic TAG in taking each lexical entry to be associated with the set of elementary tree structures that specify its local syntactic configurations. These structures can be combined via either *substitution* or *adjoining*, to produce a complete sentential analysis.

The elementary trees of D-LTAG are anchored (by and large) by *discourse connectives*, whose substitution sites can be filled by the representation of anything interpretable as an *abstract object* (ie, a proposition, fact or event). Elements so interpretable include discourse segments, sentences, clauses, nominalisations and demonstrative pronouns.

D-LTAG structures are interpreted in terms of discourse relations between arguments, using a combination of compositional semantics, inference and anaphor resolution. The latter is essential for interpreting discourse adverbials, which convey a discourse relation between the *abstract object* (AO) interpretation of the matrix clause and an AO interpretation in the discourse

context. That discourse adverbials differ in this way from structural connectives is demonstrated on theoretical grounds in Webber *et al.* (2003) and on empirical grounds in Creswell *et al.* (2002). An explanation for the anaphoric character of discourse adverbials is given in Forbes (2003) and Forbes-Riley *et al.* (2006), along with details on D-LTAG's syntactic-semantic interface.

Note that, although D-LTAG produces only analyses in the form of trees, it is acknowledged in Webber *et al.* (2003)- as noted earlier by Bateman (1999) and Gardent (1997)- that a discourse unit must be allowed to participate in one constituent structure with left-adjacent material and another with right-adjacent material. This would require relaxing substitution constraints in D-LTAG that the same tree only substitutes into a single site.

Empirical data on the predicate-argument structure of discourse connectives are now available in Release 1.0 of the annotated Penn Discourse TreeBank² (Dinesh *et al.*, 2005; Miltsakaki *et al.*, 2004a,b; Prasad *et al.*, 2004; PDTB-Group, 2006; Webber, 2005). In the PDTB, discourse connectives have been manually annotated, along with the text spans that give rise to their two arguments. The text that has been annotated is the same 1-million word Wall Street Journalcorpus that has been annotated for syntactic structure in the Penn TreeBank and for verbs and their arguments in PropBank³. Annotated in Release 1.0 are tokens of the following explicit connectives and their arguments:

- 31 subordinating conjunctions (eg. *when, because, as soon as, now that, etc.*) and their modi_ers (*only, just, even, mainly, etc.*)

² <http://www.seas.upenn.edu/~pdtb>

³ <http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T14>

- 7 coordinating conjunctions (*and, but, or, either/or, nor, neither/nor, so*)
- 62 discourse adverbials (*eg., otherwise, meanwhile, consequently, for example, etc.*)

This constitutes 18505 tokens, one for each instance of an explicit connective and its arguments. Also annotated in Release 1.0 are implicit connectives in three sections of the WSJ corpus, comprising 2003 tokens. An early effort to use data in the PDTB to develop a procedure for resolving the anaphoric argument of the discourse adverbial *instead* is described in Miltsakaki *et al.* (2003), and the effect of *Information Structure* on the preferred argument of the discourse adverbial *otherwise* is described in Kruijff-Korbayová and Webber (2001). Although English is currently the only language for which there is a Discourse TreeBank, the idea of a lexicalized discourse-level grammar and a discourse treebank should be of interest and of value for other languages as well.

References

- Bateman, J. (1999). The dynamics of ‘surfacing’: An initial exploration. In Proceedings of International Workshop on Levels of Representation in Discourse (LORID’99), pages 127-133, Edinburgh.
- Creswell, C., Forbes, K., Miltsakaki, E., Prasad, R., Joshi, A., and Webber, B. (2002). The discourse anaphoric properties of connectives. In Proceedings of the Discourse Anaphora and Anaphor Resolution Colloquium, Lisbon, Portugal.
- Dinesh, N., Lee, A., Miltsakaki, E., Prasad, R., Joshi, A., and Webber, B. (2005). Attribution and the (non-)alignment of syntactic and discourse arguments of connectives. In ACL Workshop on Frontiers in Corpus Annotation, Ann Arbor MI.
- Forbes, K. (2003). Discourse Semantics of S-Modifying Adverbials. Ph.D. thesis, Department of Linguistics, University of Pennsylvania.

-
- Forbes-Riley, K., Webber, B., and Joshi, A. (2006). Computing discourse semantics: The predicate-argument semantics of discourse connectives in D-LTAG. *Journal of Semantics*, 23, 55-106.
- Gardent, C. (1997). Discourse tree adjoining grammars. Claus report nr.89, University of the Saarland, Saarbrücken.
- Kruijff-Korbayová, I. and Webber, B. (2001). Information structure and the semantics of “otherwise” In *ESLLI’2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*, pages 61-78, Helsinki, Finland.
- Miltsakaki, E., Creswell, C., Forbes, K., Prasad, R., Joshi, A., and Webber, B. (2003). Anaphoric arguments of discourse connectives: Semantic properties of antecedents versus non-antecedents. In *EACL Workshop on Computational Treatment of Anaphora*, Budapest, Hungary.
- Miltsakaki, E., Prasad, R., Joshi, A., and Webber, B. (2004a). Annotating discourse connectives and their arguments. In *NAACL/HLT Workshop on Frontiers in Corpus Annotation*, Boston.
- Miltsakaki, E., Prasad, R., Joshi, A., and Webber, B. (2004b). The penn discourse treebank. In *LREC, Lisbon, Portugal*. PDTB-Group, T. (2006). *The Penn Discourse TreeBank 1.0 annotation manual*. Technical Report IRCS 06-01, University of Pennsylvania.
- Prasad, R., Miltsakaki, E., Joshi, A., and Webber, B. (2004). Annotation and data mining of the Penn Discourse TreeBank. In *ACL Workshop on Discourse Annotation*, Barcelona, Spain.
- Schabes, Y. (1990). *Mathematical and Computational Aspects of Lexicalized Grammars*. Ph.D. thesis, Department of Computer and Information Science, University of Pennsylvania.
- Webber, B. (2005). A short introduction to the Penn Discourse TreeBank. In *Copenhagen Working Papers in Language and Speech Processing*.
- Webber, B., Stone, M., Joshi, A., and Knott, A. (2003). Anaphora and discourse structure. *Computational Linguistics*, 29, 545-587.

How Information Structure Shapes Prosody

Caroline Féry *Shinichiro Ishihara*

University of Potsdam

caroline.fery@googlemail.com *s_i@alum.mit.edu*

1 Introduction

In this talk, we discuss interaction of prosodic phrasing, register scaling, and tonal phenomena, based on various experimental results. We propose that information structure (IS) affects register scaling rather than prosodic phrasing. We also show that in addition to prosodic phrasing and IS-related effects, other tonal phenomena also need to be taken into account to interpret our experimental results correctly.

2 Phrasing

Nearly all accounts on syntax-prosody mapping (e.g., Nespor & Vogel 1986, Cinque 1993, Selkirk 1995) predict that a subject and an adjunct behave differently from an object, in forming a separate prosodic phrase from an adjacent verb. For German, Gussenhoven's SAAR has been influential, which also takes into account the information structural status of constituents. Crucially, a verb is integrated into the prosodic phrase of an adjacent argument's prosodic phrase, and the shared prosodic phrase is headed by a unique pitch accent on the argument. When an adjunct separates them, integration is blocked. When the adjunct is given, however, integration is possible again.

Our experiments show that integration in German is a gradient phenomenon. Considering the presence of a pitch accent as an indicator of a prosodic phrase, we observe that a preceding object is nearly always phrased together with the

verb, but a subject does the same in 56% of the case. Contrary to the expectation from the previous theories, an adjunct is integrated in the preceding phrase in nearly 80%. Furthermore, focusing an adjunct does not change too much the integrational behavior of the verb, indicating that verb integration takes place regardless of the focus status of the preceding adjunct.

We propose that prosodic phrasing is in fact part of the syntax (see Ishihara 2003, Arregi 2004 and Wagner 2005 for instance), and that the realization of pitch accents is the result of a different component of prosody altogether.

3 Tonal matters

As shown by different researchers (Bruce 1977, van den Berg et al. 1992), all related prosodic phrases and intonation phrases have their own reference line, which is calculated from their location in a sentence relatively to the other phrases. In addition, their information status affects the register scaling.

Ishihara (2003) shows experimentally that Major Phrase (MaP) and what he calls Focus Intonation (FI) domain are independent. He claims that the former is purely syntax-based, while the latter is created by changing the pitch register of the focused phrase (by F0-boosting) and the post-focal material (by F0-compression). His experimental results show that downstep, which lowers the pitch realization of non-initial pitch accents within a MaP, and the post-focal compression, which lowers the pitch realization of post-focal material, are independent phenomena. Pitch reset after downstep at the following MaP boundary, and pitch reset after an FI are observed independently. Furthermore, FI can be embedded, while it has been assumed that embedding of prosodic phrases such as MaP is not allowed under the Strict Layer Hypothesis.

Also in German, scaling of accents is directly influenced by focus and givenness. In Féry & Ishihara (2005) and Ishihara & Féry (2006), we have

shown that the pitch height of a prenuclear word varies, depending on its status as First Occurrence Focus (a focused expression that appears in the discourse for the first time), Second Occurrence Focus (a focused expression that has already appeared in the discourse and is repeated) or non-focused. Focus boosts the reference line of a prosodic phrase, and givenness lowers it. We attribute this effect to the influence of information structure on whole prosodic phrases. We also show that other independent phonological effects must be taken into account. Postnuclear elements, for example, cannot realize any pitch accents regardless of their IS status, because postnuclearly the reference line is lowered to the speaker's bottom line.

In new experiments, it appears that the reference line scaling is not enough to compute the real value of the individual tones in German, but purely tonal effects, like H-raising, downstep and H-lowering, are also active. Beside downstep, H-raising is the effect of a final L tone on the preceding H tone, and H-lowering is the effect of a raised H-tone on a preceding H-tone.

4 Model

Tonal effects are pervasive in the reflexes of information structure, and the relationship between individual accents in a sentence is playing a very important role. An efficient model of prosody has to consider these partial influences on the pitch accents and bring them together. Syntax influences phrasing, and information structure influences tonal scaling. The presence of pitch accent is just one aspect of prosody, and how the pitch accents interact is at least as important as their mere presence.

Metrical grids and a rule like Gussenhoven's SAAR are not able to represent pitch scaling. Much better are metrical trees, standing for both register and

tones. Prosodic phrasing is an independent matter, which is best located in the syntactic component.

References

- Berg, Rob van de, Carlos Gussenhoven & Toni Rietveld (1992) Downstep in Dutch: Implications for a model. In Gerhard Docherty & D. Robert Ladd (eds.) *Papers in Laboratory Phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 335-367.
- Bruce, G. (1977) Swedish word accent in sentence perspective. *Travaux de l'Institut de Linguistique de Lund* 12, Gleerup, Lund.
- Cinque, Guglielmo (1993) A null theory of phrase and compound stress. *Linguistic Inquiry* 24.2: 239-297.
- Féry, Caroline & Shinichiro Ishihara (2005) *Interpreting Second Occurrence Focus*. Ms. Potsdam.
- Gussenhoven, Carlos (1992) Sentence accents and argument structure. In Iggy Roca (Ed.) *Thematic Structure. Its role in grammar*. Foris. Berlin. 79-106.
- Ishihara, Shinichiro (2003) *Intonation and interface conditions*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Ishihara, Shinichiro, Caroline Féry & Gisbert Fanselow (in prep) *Integrational effects in German, Japanese and Hungarian*. Ms. Potsdam.
- Nespor, Marina & Irene Vogel (1986) *Prosodic phonology*. Dordrecht: Foris.
- Selkirk, Elisabeth O. (1995) *Sentence Prosody: Intonation, Stress and Phrasing*. In: Goldsmith, John (ed.) *Handbook of Phonological Theory*. Blackwell. 550-569.

Autonomy of Syntax and Information Structure

Gisbert Fanselow

University of Potsdam

fanselow@uni-potsdam.de

1 The issue

Information Structure is reflected in the choice of syntactic constructions in many if not all languages. One class of models (represented, e.g., by Rizzi 1997) assumes that there is a direct and transparent mapping between syntax and information structure, such that features related to information structure are encoded in the syntax. A different class of models (see, e.g., Chomsky 2005) subscribes to the view that syntax is autonomous, such that, in particular, no features that are not properties of items in the lexicon can figure in syntactic computations. For many (if not all) languages this implies that concepts of information structure play no role in syntactic derivations. Syntax interacts only in an indirect way with information structure: the relation is mediated by the syntax-PF interface. The second type of model can be shown to be empirically superior for at least a certain class of languages.

2 Is syntax sensitive to focus?

If syntax would represent information structure in a direct and transparent way, we would expect that there are structural slots reserved for focus phrases, or that focus plays a role in triggering movement. But what appears to be a focus position often (if not always) turns out not to be one under closer inspection. Thus, in German, Turkish, Hindi, and further SOV languages, the preverbal ‘focus’ position is obligatory for focused objects only. Focussed subjects and

adjuncts can, but need not appear there. The best analysis for these facts involves a prosodic alignment principle for stressed XPs, with a preference for right-alignment in IP. Focussed objects already are rightmost (and have no reason to move away). Subjects and adjuncts do not move to the right, rather, de-stressed objects can be scrambled across subjects/ adjuncts, in which case the latter elements appear preverbally. Given the optional nature of scrambling, we understand why stressed non-objects need not appear in preverbal position. There is no preverbal focus position, but rather a preference for right-aligned stressed XPs.

Likewise, the left-peripheral position in CP to which focussed phrases may move in several languages (Hungarian, German, Czech, Russian, Italian, Greek) is not a focus position. Evidence from the movement of idioms and intervention effects (see Fanselow & Lenertova 2006) reveal that the crucial properties of the movement element is its bearing stress, which can, but need not be related to being in focus.

The well-formedness of syntactic representations and the licensing of movement operations is therefore often linked to stress, -and not to focus- in quite a number of languages. Since stress is related in a complex way to focus, one may sometimes get the incorrect impression that syntax is sensitive to focus, too.

3 A Model for focus-free syntax

Autonomous “narrow” syntax itself would still have a hard time being sensitive to factors such as stress, however! Stress is not a lexical feature and movement is triggered by (abstract) features of lexical items (Chomsky 1995, 2005). The impact of stress on syntax must thus be explained differently. In narrow syntax, XPs move independent of their prosodic properties. This leads to overgeneration

in contexts in which, e.g., only the leftmost stressed XP can move. Structures with an incorrect placement of stressed XPs are filtered out when syntactic structures are interpreted prosodically. Scrambling (usually displacing unaccented material) and movement to the left periphery (observing a superiority-like constraint) do not change prominence relations among stressed elements. This restriction of what movement does (it must not change prominence) can be expressed in terms of the copy & deletion theory of movement: the copy of the moved phrase in the target position can be spelled out there only if this does not imply a reversal of the prominence relations established so far.

4 An apparent problem explained away

If focus is a notion formally interpreted by prosody only, how can it have an impact on interpretation in a Y-/T-like model of grammar, in which there is no direct link between PF and LF, such that prosodic distinctions can be interpreted at LF only if they are reflected in the syntax? The answer to this question fails to motivate the use of focus features in the syntax, however: that a phrase is in focus is not so much a prosodically determined feature but a property of the context of an utterance. “Focality” is anchored in the context. Context influences interpretation, and whatever impact focus may have on meaning can be described in the context-meaning interface alone. Certain prosodic contours are, however, inappropriate in certain contexts, and interpretations are interpretations of appropriate utterances only. There is no need for formal links for an exchange of information between PF and LF. All we need is context.

5 Further issues

Possible extensions of this line of research to topicality (in German) can be found in Fanselow (2006). In a weak interpretation of our model, we confine the claim that information structure is irrelevant for syntax to those languages in which it has no morphological reflex (in the form of topic or focus markers, or corresponding agreement morphology), this being linked to the claim that languages may differ with respect to the lexical features they employ. A stronger interpretation claims that syntax never responds to information structure in a direct way. The obvious task then lies in reanalysing focus- and topic markers as elements of a different sort.

Narrow syntax must be complemented by a model of “construction” choice. The same “construction” (say, passive) may have quite different functions (conservation of person-number-hierarchies, information structure, suppression of agent) in different languages.

References

- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT-Press.
- Chomsky, N. 2005. On phases. Ms.
- Fanselow, G. 2006. On Pure Syntax (Uncontaminated by Information Structure). Ms.
- Fanselow, G. & D. Lenertova. 2006. Left Peripheral Focus: Mismatches between Syntax and Information Structure. Ms., Universities of Potsdam and Leipzig.
- Rizzi L. 1997. The Fine Structure of the Left Periphery. In: L. Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, 281-337

Quantification and Topicality

Cornelia Endriss Stefan Hinterwimmer

University of Potsdam Humboldt University of Berlin

endriss@uni-potsdam.de stefan.hinterwimmer@rz.hu-berlin.de

In this talk, we aim at reducing two seemingly unrelated phenomena to one underlying principle: the fact that in adverbially quantified sentences topic marked indefinites are interpreted in the restrictor of the Q-adverb, and the fact that in sentences with two or more quantificational DPs topical indefinites take widest scope.

It is widely assumed that syntax determines the arguments of determiner quantifiers (D-quantifiers), whereas information structure (IS) is the decisive factor in the definition of restrictor and nucleus in constructions with adverbial quantifiers (A-quantifiers) (cf. Rooth, 1985 and Partee, 1991 among many others).

Despite many differences, the theories that explain the different interpretations occurring with A-quantifiers as an information structural effect agree that topical/non-focal material is mapped onto the restrictor and focal/non-topical material is mapped onto the nuclear scope (cf. (1a) and (2a) from Krifka, 2001 and their interpretations and intuitive meanings in (1b, c) and (2b, c), respectively).

- (1)
 - a. A freshman usually wears a BASEball cap.
 - b. MOST s . $[\exists x. \text{freshman}(x) \wedge \text{in}(s, x)]$
 $[\exists x. \text{freshman}(x) \wedge \text{wears_bc}(s, x)]$
 - c. ‘Most freshmen are such that they wear a baseball cap.’
- (2)
 - a. A FRESHman usually wears a baseball cap.
 - b. MOST s . $[\exists x. \text{wears_bc}(s, x)] [\exists x. \text{freshman}(x) \wedge \text{wears_bc}(s, x)]$
 - c. ‘Most baseball cap wearers are freshmen.’

We can hence formulate the following observation:

Observation 1 (cf. e.g. Partee 1991): Topical material tends to be interpreted in the restrictor of an adverbial quantifier.

Concerning D-quantification, it has been observed that the topical status of a quantificational DP also affects its interpretation. Whereas restrictor and nucleus of a D-quantifier are determined widely independent of information structure, the topical status of a quantificational DP still contributes to the truth conditions of the sentence: a topical DP can only receive a strong interpretation and can either be interpreted generically or 'specifically', i.e. as taking wide scope over all other involved operators. The generic interpretation is shown in (3a), whereas the contrast in (4a) vs. (4b) illustrates the wide scope interpretation, which is often referred to as a 'specific' interpretation ('RP' stands for 'resumptive pronoun').

- (3) a. Ein Löwe, der hat eine lange Mähne.
 A lion RP has a long mane
 'Lions have long manes.'
- b. GEN s. $[\exists x. \text{lion}(x) \wedge \text{in}(s, x)] [\exists x. \text{lion}(x) \wedge \text{has_long_mane}(s, x)]$
- (4) a. Einen Löwen, den hat jeder gesehen.
 A lion RP has everybody seen.
 'There is a particular lion that has been seen by everyone.'
- b. Einen Löwen hat jeder gesehen.
 A lion has everybody seen.
 'Everyone has seen a (particular) lion.'

Both strong interpretations are induced by German left dislocation, which marks the left dislocated element unambiguously as topic (see Frey, 2004). The following observation has been made on grounds of sentences such as (4):

Observation 2 (cf. e.g. Cresti 1995): Topical material tends to take wide scope.

The interpretation of a topical DP as generic can easily be explained as an instance of *Observation 1*. If it is assumed that a covert Q-adverb with generic force is inserted in sentences such as (3a), the result will be an interpretation as is given in (3b). However, it is not at all obvious how the two observations relate to each other. In our talk, we want to provide an answer to this question. We will argue that both observations result from one and the same principle, which is given below.

- (5) *Topic Occurrence Principle:* Topical material cannot be interpreted in the nuclear scope of a quantifier.

The *Topic Occurrence Principle* suggests itself if quantification is understood as a higher order predication process where the nucleus naturally corresponds to the predication of the sentence, whereas the restrictor is naturally understood as the object of predication, i.e. the topical part, while the role of the quantificational determiner itself is to specify the degree to which the predicate applies to this object (cf. Löbner 2000). The actual underlying principle would then be: ‘*Topical material resists predicative environment*’.

Observation 1 directly follows from this principle. In adverbial quantification, one option for topical material to escape the nuclear scope of the respective A-quantifier is of course to be interpreted in the restrictor, as in (1) and (2). However, there also is another – less acknowledged – possibility,

namely that the topical material is interpreted outside of the scope of the Q-adverb, thus receiving a wide scope interpretation (cf. (6)).

- (6) a. Ein Löwe, der ist meistens schlecht gelaunt.
 A lion RP is usually bad-tempered
 ‘There is a particular lion that is usually bad-tempered.’
 b. $\exists x. \text{lion}(x) \wedge \text{most } s [\text{in}(s, x)] [\text{bad-tempered}(s, x)]$

With D-quantifiers, matters are different, as those quantifiers choose their arguments syntactically. This means that topical material that does not belong to the syntactic complement of a D-quantifier cannot end up in its (semantic) restrictor. It also cannot be interpreted in the nuclear scope due to the *Topic Occurrence Principle*. So it has to be interpreted outside of the scope of the D-quantifier. Hence, the only option for a topical DP is to take wide scope. The other D-quantifier accordingly has to be interpreted in the nuclear scope of the topical one. This explains the contrast in (4).

Generally speaking, the configuration in (7a) allows for four different imaginable interpretations, which are given in (7b). The *Topic Occurrence Principle* in combination with the fact that D-quantifiers are IS-insensitive now explains that only the reading in 3. is a viable option.

- (7) a. Surface Structure: ... DP₁ (...) (... (DP₂)_{Top} ...)
 b. Imaginable Interpretations:
 1. Q₁ [...]_{Restr} [... Q₂ ...]_{Nucl}
 2. Q₁ [... Q₂ ...]_{Restr} [...]_{Nucl}
 3. Q₂ [...]_{Restr} [... Q₁ ...]_{Nucl}
 4. Q₁ [... Q₂ ...]_{Restr} [...]_{Nucl}

References

- Cresti, Diana (1995): Indefinite Topics. Dissertation, MIT.
- Frey, Werner (2004): Notes on the Syntax and the Pragmatics of German Left Dislocation, in: H. Lohnstein & S. Trissler (eds.), *The Syntax and Semantics of the Left Periphery*, Berlin: Mouton de Gruyter, 203-233.
- Krifka, Manfred (2001): Non-novel Indefinites in Adverbial Quantification, in: C. Condoravdi & G. Renardel de Lavalette, *Logical Perspectives on Language and Information*, CSLI Publications, Stanford 2001, 1-40.
- Löbner, Sebastian (2000): Polarity in Natural Language: Predication, Quantification and Negation in Particular and Characterizing sentences, *Linguistics and Philosophy* 23: 213-308.
- Partee, Barbara (1991): Topic, Focus and Quantification, in: S. Moore & A. Z. Wyner (eds.), *Proceedings from Semantics and Linguistic Theory*, volume 1 of Cornell Working Papers in Linguistics.
- Rooth, Mats (1985): *Association with Focus*. Ph.D. dissertation, University of Massachusetts at Amherst

Interpreting Topics in Quantificational Structures

Cornelia Endriss Stefan Hinterwimmer

Universität Potsdam Humboldt Universität, Berlin

endriss@uni-potsdam.de stefan.hinterwimmer@rz.hu-berlin.de

Our poster consists of three parts, which represent three different aspects of our project work. In the first section we summarize our work on adverbial quantification, in the second section we summarize our work on determiner quantification, and in the third section we summarize the main features of a proposal that aims at reducing the phenomena observed in the two domains to one underlying principle.

1 Quantificational adverbs and topicality

In this part of the project, the following two questions are of central importance:

- (1) How do information structural categories like topicality and focality determine exactly which part of a given clause is interpreted in the restrictor of an adverbial quantifier, and which part is interpreted in the nuclear scope (see e.g. Partee 1991, Rooth, 1995)?
- (2) Are Q-adverbs only able to quantify over situations/eventualities (see von Stechow 1994), or are they also (in light of the phenomenon known as *Quantificational Variability*) able to quantify over individuals (see Kamp 1981 and Heim 1982)?

Concerning the first question, a comparison between adverbially quantified sentences with indefinites and ones with universally quantified DPs (in German and English) shows that the role of information structure with respect to clause splitting is an indirect one: information structure has an influence on the relative order of the Q-adverb and the DPs contained within a given clause, which in turn plays a role in the compositional interpretation of the clause. Furthermore, information structure helps resolving the free variable over situation predicates

that always forms a part of the restrictor of a Q-adverb (see von Stechow 1994). On the other hand, there is no mapping algorithm which has direct access to information like focus marking and/or topic marking (cf. Beaver and Clark 2003), and is furthermore able to overwrite syntactic relations like c-command. Based on these assumptions, detailed analyses were developed that account for the conditions under which sentences containing the above mentioned types of DPs allow co-variation with the situations quantified over by the respective Q-adverb (Hinterwimmer 2005, Hinterwimmer in preparation).

Concerning the second question, we show (Endriss and Hinterwimmer to appear-a, Hinterwimmer 2005) that strong evidence supports the assumption that Q-adverbs are only able to quantify over situations/eventualities. This evidence is based on the following observation: adverbially quantified sentences that contain (de-accented) indefinites modified by relative clauses only show *Quantificational Variability Effects* (QVEs) if the tense of the matrix verb agrees with the tense of the relative clause verb. On the other hand, no comparable restriction is observed in the case of sentences that contain quantificational DPs modified by relative clauses – which is unexpected, if Q-adverbs have (at least potentially) the same quantificational domain as quantificational determiners.

Our explanation for this phenomenon is based on the assumption that Q-adverbs quantify exclusively over situations, and that these situations have to be located within an interval that is determined on the basis of locally available information. This has the consequence that the situations quantified over have to be located within the running times of the relative clause situations – which results in a necessary contradiction if the matrix verb is marked for past tense, while the relative clause situations are located in the past.

For reasons of space, this part of our work on adverbial quantification is not dealt with on the poster.

2 Quantificational determiners and topicality

In this section of the poster we summarize our work on the interaction of information structure with the scope-taking behaviour of various types of quantifiers. An analysis has been developed which reduces the fact that certain quantificational DPs take exceptional, island-violating wide scope to the topical status of these DPs (Ebert and Endriss 2004, Endriss in preparation). This account furthermore allows for a natural classification of quantificational DPs into ones that can become topics, and ones that can't. This classification is exclusively based on the semantic properties of the respective quantifiers. In Endriss (in preparation) it is furthermore shown that exceptional wide-scope is a real phenomenon that cannot be reduced to pragmatic mechanisms like speaker reference (Fodor and Sag 1982) or domain restriction (Schwarzschild 2002).

We allow for functional topics also (such as '*some picture of himself*') and predict a functional wide scope reading for them. This reading can be distinguished empirically from an ordinary narrow scope reading by way of the pair-list/natural function differentiation known from the semantic investigations of questions. An ordinary narrow scope indefinite in the scope of e.g. a universal quantifier allows for a pair-list elaboration, whereas a functional wide scope indefinite supports only functional elaborations.

-
- (1) a. A freshman usually wears a BASEball cap.
 b. MOST s . $[\exists x. \text{freshman}(x) \wedge \text{in}(s, x)]$
 $[\exists x. \text{freshman}(x) \wedge \text{wears_bc}(s, x)]$
 c. ‘Most freshmen are such that they wear a baseball cap.’
- (2) a. A FRESHman usually wears a baseball cap.
 b. MOST s . $[\exists x. \text{wears_bc}(s, x)] [\exists x. \text{freshman}(x) \wedge \text{wears_bc}(s, x)]$
 c. ‘Most baseball cap wearers are freshmen.’

We show that all types of DPs can in principle receive functional interpretations (contra Winter 2004). It is thus predicted that all those indefinites that take exceptional wide scope also allow *exceptional wide functional* readings, as the topic interpretation mechanism that we propose allows the computation of functional as well as non-functional DPs. Functional as well as non-functional wide-scope interpretations can thus be accounted for via the same mechanism, which in effect makes choice-function approaches to wide scope phenomena superfluous. All the more so as these have been shown to be unable to account for certain readings where an indefinite takes island-violating wide scope (cf. Schwarz 2001).

3 Semantic effects of topicality with quantificational determiners and adverbs

In this section of the poster we show that the phenomena dealt with in the other two sections can be reduced to one underlying principle: topical DPs cannot be interpreted in the nuclear scope of another quantifier. This part of our work is dealt with in detail in our talk *Quantification and Topicality*.

References

- Beaver, David and Brady Clark (2003): *Always and Only: Why not all Focus-Sensitive Operators are Alike*. *Natural Language Semantics* 11, 323–362.
- Cresti, Diana (1995): *Indefinite Topics*. PhD thesis, MIT.
- Ebert, Christian & Cornelia Endriss (2004): Topic Interpretation and Wide Scope Indefinites, in: K. Moulton & M. Wolf (eds.), *Proceedings of NELS 34*, Amherst: GLSA.
- Endriss, Cornelia (in preparation): Quantificational Topics. A non-functional treatment of wide scope phenomena, PhD thesis.

Anaphora Resolution in Discourse: The Role of Pitch Range and Pause

Ekaterina Jasinskaja Ulrike Kölsch Jörg Mayer

University of Potsdam

{jasinsk,ukoelsch,jmayer}@uni-potsdam.de

This paper presents the results of a study on the influence of prosody on the interpretation of anaphoric pronouns. Until now, empirical studies in this area have predominantly concentrated on local prosodic features of pronouns, such as pitch accent (cf. e.g. Venditti et al., 2002), whereas the impact of global prosodic parameters of an utterance, such as pitch range or pause duration, was almost entirely ignored. At the same time, the current state of research on discourse anaphora on the one hand and global prosody on the other strongly suggests that a link must exist between these two (seemingly unrelated) phenomena.

On the one hand, it is an established fact that discourse structure affects accessibility/salience of possible antecedents to anaphoric expressions (Grosz & Sidner, 1986; Polanyi, 1988; Asher & Lascarides, 2003). One of the factors that constrain the resolution of an anaphoric pronoun is the hierarchical discourse structure of the context in which the potential antecedent is mentioned. Roughly, if a new referent is introduced in a subordinated discourse unit (subtopic) it is no longer accessible for anaphoric reference after a shift from the subtopic back to the main topic.

On the other hand, pitch range and pauses have been shown to signal the structure of a spoken monologue. Most studies agree that expanded pitch range correlates with the introduction of new discourse topics and sub-topics or with the beginning of a paragraph; compressed pitch range, on the other hand, signals

the end of a paragraph or the closing of a (sub-) topic. These results were already obtainable on the basis of a rather simplistic pre-theoretical notion of discourse structure, equating the latter either with the structure of a written text, i.e. the paragraph structure (Lehiste, 1975; Sluijter & Terken, 1993) or with a discourse topic model adopted for the specific material of the study (Ayers, 1994; Venditti & Swerts, 1996). Other studies that based their analyses on more elaborate, theoretically motivated hierarchical notions of discourse structure, such as RST or SDRT, have also shown that the width and position of the pitch range correlate significantly with the depth of embedding of discourse units (Möhler & Mayer, 2001; den Ouden et al., 2002). Similar results are reported for the duration of silent pauses. Pauses are longer before units introducing new discourse topics. The shortest pauses appear between intonational phrases dealing with the same topic (e.g. Grosz & Hirschberg, 1992; Swerts, 1997).

However, it is still largely an open question whether hearers actually use the information encoded by pauses and pitch range to disambiguate anaphoric references, which also bears on a more general theoretical issue, whether global prosodic parameters contribute to the *linguistic* interpretation of an utterance. With the present study we want to fill this gap in empirical research. The purpose of our experiments is to test the hypothesis that global prosodic features contribute to the interpretation of linguistic expressions by disambiguating structurally ambiguous discourses. Among all the interpretation phenomena sensitive to discourse structure, we confine our attention to pronominal anaphora, which is a paradigmatic case among various types of anaphora as well as a central phenomenon of discourse interpretation.

We conducted a series of perception experiments with ambiguous discourses. In experiment 1, the pitch range of the sentences and the pause duration between sentences was manipulated. In experiment 2, only pitch range

was manipulated while all pauses were set to a standard duration. In experiment 3 the pitch range of all sentences was set to a standard interval, pauses between sentences were manipulated according to the discourse structure. The results of our experiments corroborate our main research hypothesis that global prosodic parameters such as pitch range and pause duration influence the resolution of anaphoric pronouns. But experiments 2 and 3 revealed apart from this, that only the combination of both prosodic parameters has enough strength in order to disambiguate structurally ambiguous discourses. The direction of the observed effect in experiment 1 is clearly in accordance with the predictions of the existing theories of discourse anaphora and the current state of research on discourse prosody. This result strongly suggests that, indeed, the relationship between global prosody and the choice of the antecedent of an anaphoric pronoun is mediated by the choice of attachment site of an utterance in the discourse structure.

References

- Asher, N. & Lascarides, A. (2003) *Logics of Conversation*. Cambridge University Press.
- Ayers, G.M. (1994) Discourse functions of pitch range in spontaneous and read speech. *Ohio State University Working Papers in Linguistics* 44, 1-49.
- Grosz, B. & Hirschberg, J. (1992) Some intonational characteristics of discourse structure. *Proceedings ICSLP, Banff*, 429-432.
- Grosz, B.J. & Sidner, C.L. (1986) Attention, intentions and the structure of discourse. *Computational Linguistics* 12, 175-204.
- Lehiste, I. (1975) The phonetic structure of paragraphs. In: Cohen, A. and Nöteboom, S.G. (eds.) *Structure and Process in Speech Perception*. Springer, 195-203.

- Möhler, G. & Mayer, J. (2001) A discourse model for pitch-range control. Proceedings ISCA Tutorial and Research Workshop on Speech Synthesis, Perthshire.
- den Ouden, H., Noordman, L. & Terken, J. (2002) The prosodic realization of organizational features of texts. Proceedings International Conference on Speech Prosody, Aix-en-Provence, 543-546.
- Polanyi, L. (1988) A formal model of the structure of discourse. *Journal of Pragmatics* 12, 601-638.
- Sluijter, A.M.C. & Terken, J.M.B. (1993) Beyond sentence prosody: Paragraph intonation in Dutch. *Phonetica* 50, 180-188.
- Swerts, M. (1997) Prosodic features at discourse boundaries of different strength. *Journal of the Acoustical Society of America* 101, 514-521.
- Venditti, J.J., Stone, M., Nanda, P. & Tepper, P. (2002) Discourse constraints on the interpretation of nuclear-accented pronouns. Proceedings International Conference on Speech Prosody, Aix-en-Provence, 675-678.
- Venditti, J. & Swerts, M. (1996) Prosodic cues to discourse structure in Japanese. Proceedings ICSLP, Philadelphia, 725- 728

Pragmatics and Prosody of Implicit Discourse Relations: The Case of Restatement

Ekaterina Jasinskaja

Department of Linguistics
University of Potsdam, Potsdam, Germany

jasinsk@uni-potsdam.de

The focus of this study is on implicit discourse relations-intersentential semantic relations between events, propositions, or speech acts that are not explicitly signalled by any linguistic markers, such as *then* (for temporal succession) or *because* (for a causal relation), but nevertheless are successfully inferred by the hearer. Special attention will be paid to the relation *Restatement* understood broadly as one that holds between two sentences where the second sentence redescribes the same event, or restates the same fact as that presented in the first sentence. For instance, the most natural interpretation of (1) is that Alena lost her main transportation means *in that* she broke her skis, i.e. the event of Alena's breaking her skis is identical with the event of losing her main transportation means.

- (1) Alena broke her skis.
She lost her main transportation means.

The main questions that this study seeks to answer is how and why implicit discourse relations are inferred. There are two major positions in pragmatics concerning this issue. According to one approach, most prominently represented by Segmented Discourse Representation Theory (SDRT, Asher & Lascarides, 2003), a discourse must be *coherent*, and making sense of a discourse boils down to figuring out in which way it fulfils this requirement. The other position

is that a discourse must first of all be *relevant*, i.e. fulfil its communicative goal, so coherence is not viewed as a goal in itself, but a by-product of figuring out in which way a discourse is relevant. This is an appealing approach since it tries to explain coherence instead of simply postulating it and ultimately derive discourse relations from the general principles of rational and cooperative behaviour. However, the formalisation of these ideas (Grice, Neo-Griceans, Relevance Theory, etc.) is only starting to be developed. The purpose of this study is to make a contribution to formalising the inference of discourse relations based on Gricean principles.

In this context particularly interesting insights can be gained by considering discourse connectives like *and* (2a) and prosody, e.g. utterance-final ‘continuation rise’ (↗), cf. (2b). Both expressive devices have rather abstract, almost empty semantics, so it is not very plausible to assume that they signal discourse relations in a more or less direct way. Nevertheless, they do affect the inference of discourse relations. Thus the *Restatement* interpretation is not as prominent in (2a) and (2b) as it is in (1), a different reading becomes more accessible: Alena broke her skis, and *in addition to that* she lost her main transportation means (e.g. her car was stolen). The connections between the surface form and the inferred relations must be very indirect in these cases, possibly by-passing a whole number of different mechanisms involved in discourse interpretation. Uncovering these connections is therefore likely to shed light on the mechanisms involved, and ultimately, on the questions how and why implicit discourse relations are inferred.

(2) a. Alena broke her skis,
and (she) lost her main transportation means.

b. Alena broke her skis (↗),
she lost her main transportation means (↗)

<http://www.sfb632.uni-potsdam.de>

The theory developed in the present study is cast in the framework of dynamic update semantics (Dekker, 1993) enriched with the notion of exhaustive update largely adopted from the work of van Rooij and Schulz (2004) and the principles of discourse topic management related to the ideas of Zeevat (2005). In the proposed theory the interpretation of a discourse is a sequence of exhaustive and non-exhaustive updates of the initial information state s with the meanings of individual utterances, cf. (3). The update function, in particular the exhaustive update, depends on the discourse topic, or the question under discussion (QUD). Thus, roughly, if the current QUD is *What happened?*, the topic is the question predicate *happen* (derived from the interrogative by abstracting over the *wh*-phrases), and the exhaustive interpretation of the sentence *Alena broke her skis* says that Alena breaking her skis is the *only* (relevant) event that happened.

- (3) $s \parallel \exists e1 [A. \text{ broke skis } (e1) \wedge \text{ happen } (e1)] \parallel \parallel \text{happen} \parallel \exists e2 [A. \text{ lost transport. means } (e2) \wedge \text{ happen } (e2)] \parallel \parallel \text{happen}$

The central role in the inference of *Restatement* is played by two default principles: the Principle of Exhaustive Interpretation, and the Principle of Topic Continuity. The former says that by default, an utterance undergoes exhaustive update; the latter requires that the discourse topic not change unless explicitly signalled otherwise. If both sentences in (1) are exhaustivized with respect to the same question, e.g. *What happened?*, we get that both Alena breaking her skis is the only event that happened, and Alena losing her main transportation means is the only event that happened, thus they must be the same event.

The function of continuation intonation in this framework consists in signalling non-exhaustivity which blocks the inference of event coreference: if

Alena breaking her skis is not necessarily the only event that happened, and neither is her losing her main transportation means, nothing follows with respect to the coreference relations between these events, cf. (2b). The function of the conjunction *and* can be analysed in a number of different ways. One possibility is to follow Blakemore & Carston (1999) and assume that *and* makes one utterance out of two which only undergoes exhaustive update as a whole. The other option is to treat *and* as a signal of local topic discontinuity: the topic the current utterance is not exactly the same as that of the previous one, but stands in an additive relation to it, e.g. *What happened?*, *What else happened?*, etc. In both cases the inference of event coreference does not go through and the contrast between (1) and (2a) is correctly captured.

References

- Asher, Nicholas & Alex Lascarides (2003) *Logics of Conversation*. Cambridge University Press.
- Blakemore, Diane & Robyn Carston (1999) The Pragmatics of And-Conjunctions: The Non-Narrative Cases. UCL Working Papers in Linguistics 11, 1-20.
- Dekker, Paul (1993) *Transsentential Meditations*. PdD thesis, ILLC-Department of Philosophy, University of Amsterdam.
- Van Rooij, Robert & Katrin Schulz (2004) Pragmatic Meaning and Non-Monotonic Reasoning: The Case of Exhaustive Interpretation. Submitted to Linguistics and Philosophy.
- Zeevat, Henk (2005) Optimality Theoretic Pragmatics and Rhetorical Structure. Handout of the talk and the 6th International Tbilisi Symposium on Language, Logic and Computation.

Beyond Topic and Focus: Extraposition in Turkish

Elke Kasimir, Özgür Şahin, Juliane Böhme

University of Potsdam, Germany

elke.kasimir@catmint.de

1 Beyond topic and focus

What is the semantic-pragmatic content of the various forms of information structuring expressed by various prosodic, morphological and/or syntactic means in the languages of the world? This question does not only arise for the relatively well-known distinctions topic-comment and focus-background - the latter having been subject to extensive research in the SFB 632 / A4 project (Beck 2005, Elbourne 2005, Kasimir 2005) - but also for some less established phenomena which are related to but distinct from topic-comment, focus-background. Well-known examples from English include:

- (1) *exceptional focus projection from the subject*
 - a. The CHILD is crying. (marked)
 - b. The child is CRYing. (unmarked)
 - c. ??FATHER is crying. (usually lacks pragmatic licensing)

- (2) *PP-extraposition*
 - a. She read a BOOK yesterday about bats. (marked)
 - b. She read a book about BATS yesterday. (unmarked)
 - c. ??She lost a BOOK yesterday about bats.
(usually lacks pragmatic licensing)

In all three cases, a particular prosodic and/or syntactical configuration seems to interact with focusing/de-focusing and

topicalization/anti-topicalization; on the other hand, not well-understood pragmatic licensing conditions seem to play a decisive role.

Our empirical research into sentence-level *extraposition* in Turkish (*devrik cümle* construction) targets another such construction:

- | | | | |
|-----|--------------|---------------------------------------|-------------------------------------|
| (3) | a. Onunla | gezDIK 3-4 saat. | (marked) |
| | | <i>with-her (I) walked 3-4 hours.</i> | (unmarked) |
| | b. Onunla | 3-4 saat gezdik. | |
| | c. ??Onunlan | gezdik 12-14 saat. | (usually lacks pragmatic licensing) |

(for more examples, see (4) and (7) below). Our aim is a precise understanding of the licensing conditions underlying such data in Turkish, and an explanation of the ways in which extraposition interacts with focus-background and topic-comment.

2 Investigating extraposition in Turkish

Since Turkish is a straight-forward verb-final language, extraposition, that is the occurrence of overt material at the right of the verb, is usually easily diagnosed. More than one constituent can be moved behind the verb of the matrix sentence, as long as some still poorly understood island conditions are satisfied, and, of course, the characteristic interpretative licensing conditions on extraposition are met.

In order to find out more about the interpretative licensing conditions of extraposition, we collected, analyzed and discussed 45 real-life examples from a Turkish internet magazine *haber.com*.

Based on our work with the corpus data, we also designed a web-based questionnaire. The questionnaire presents variants of a sentence in a particular

discourse context. The participants (between 7 and 20 per item) were asked to rate the absolute and the relative acceptability of each variant.

3 Relation to focus/topic

We found that extraposed phrases cannot be *focus phrases*, but are on the other hand also often neither in any sense *given*, nor related to an implicit or explicit question (*discourse topic*). We also found that extraposed material cannot be topical in those cases where the topic of a sentence is not presupposed.

There is a general agreement that extraposed material cannot bear an accent in Turkish (see Erguvanlı 1984, Kornfilt 1997 and others). Somewhat astonishingly, we found examples where extraposition *can* be accented in highly marked contexts. Crucially, in these contexts, the extraposed phrase is still not a focus phrase, although it seems to be part of a sentence-wide focus. For an example, see (7) below.

4 Approaching the licensing conditions

We basically propose that sentences containing extraposition amount to the presentation of *two related utterance meanings*: One utterance corresponds to core sentence and excludes the extraposed material; this core sentence is interpreted as an elliptical expression. The other utterance corresponds to the whole sentence. The crucial licensing condition can then be stated as follows:

- (4) The core sentence already expresses the main communicative intention.

Example:

- (5) [Onunla gezdik] Berlin'i. (= [with-her (I) walked] in-Berlin
core sentence paraphrase: with her (I) walked somewhere

Assuming that utterance meanings expressing the same communicative intention usually agree in the choice of the sentence topic, and in the choice of the addressed discourse topic, the aforementioned interaction with information structure largely follows. As for the requirements on accent placement, we propose the following:

- (6) Extraposed material is accented iff it is involved in a contrastive relation.

Due to (4), this contrastive relation must be *pure*, that is, non-focal and non-topical:

- (7) Why did you break up? [Ben oturuyorum] Berlin'de, [o oturuyor] NewYork'ta
 (= [I live] in Berlin, [she lives] in NewYork).
core sentence paraphrase: I live somewhere, she lives somewhere else.

Notice that the core sentence can already be constructed as involving contrasting spatial locations; furthermore, as Manfred Krifka pointed out to us, the particular involved locations (NewYork, Berlin) play are not important for the communicative intention, and the resp. phrases are therefore plausibly neither focused nor topical. This seems to be crucial: (8), where NewYork and Berlin are focal, was rated much worse than (7):

- (8) Where do you to live? [Ben oturuyorum] Berlin'de, [o oturuyor] NewYork'ta
 (= [I live] in Berlin, [she lives] in NewYork).

Reference

- Beck, S. (2005). Intervention Effects Follow from Focus Interpretation. In Natural Language Semantics.
- Birtürk, A. (1998), 'A Computational Analysis of Turkish Using the Government Binding Approach', Master's thesis, Middle East Technical University.
- Elbourne, P. (2005). The Semantics of Ellipsis. In Dipper, S., Götze, M., Stede, M. (Eds.), *Interdisciplinary Studies on Information Structure (ISIS) 3* (63-110). Potsdam: Universitätsverlag Potsdam.
- Erguvanlı, E.E. (1984), *The function of word order in Turkish grammar*, University of California Press, Berkeley.
- Guéron, J. (1980), 'On the Syntax and Semantics of PP Extraposition', *Linguistic Inquiry* 11, 637-678.
- Göksel, A. (1998), Linearity, focus and the postverbal position in Turkish, in L. Johanson, ed., 'The Mainz Meeting: Proceedings of the Seventh International Conference on Turkish Linguistics', Harrassowitz Verlag, Wiesbaden, pp. 85-106.
- Göksel, A. & Özsoy, A.S. (2000), Is there a focus position in Turkish? 'Studies in Turkish Languages', Ali Göksel and Celia Kerslake.
- Kasimir, E. (2005), Question-answer test and givenness: Some question marks. In Ishihara, S., Schmitz, M., Schwarz, A. (Eds.), *Interdisciplinary Studies on Information Structure (ISIS), 3* (1-52). Potsdam, Universitätsverlag Potsdam.
- Kılıçaslan, Y. (2004), 'Syntax of Information Structure in Turkish', *Linguistics* 42(4), 717-765.
- Kornfilt, J. (1997), *Turkish (Descriptive Grammars)*, Routledge, London.
- Kornfilt, J. (1998), Rightward Movement, in L. Johanson, ed., 'The Mainz Meeting: Proceedings of the Seventh International Conference on Turkish Linguistics', Harrassowitz Verlag, Wiesbaden, pp. 107-123.
- Lambrecht, K. (2001), Dislocation, in: Haspelmath, M. et. al. *Language Typology and Language Universals. An International Handbook*, Berlin, Walter de Gruyter, pp. 1050-1078.

Özge, U. (2003), 'A Tune-based Account of Turkish Information Structure', PhD thesis, Middle East University.

Temürcü, Ç. (2001), 'Word Order Variations in Turkish: Evidence from Binding and Scope', Master's thesis, Middle East Technical University

3 The special status of focused subjects

The special status of focused subjects is conditioned by information structure: Subjects in canonical preverbal position are prototypically interpreted as topics (i.e. they are anti-focal) in the languages under discussion. Therefore, if a subject is in focus, this conflicts with its primary topical status and results in a construction divergent from the canonical one.

In our sample, there are two prominent strategies to deal with this conflict. First, several Gur languages employ topicless, i.e.thetic utterances (characterized by predicate incorporation) to realize SF. As a result, narrow SF and broad sentence focus are formally identical (Subject/Sentence Focus-Isomorphism) (e.g. Konni, (4)).

(4)a. *SF expressed by special verb tone and suffix -nÁ*

Q: Who hit Peter?

A: Mǎrỳ níǵí-nà-wà.

Mary hit-NA-3sg

‘MARY hit him.’

b. Q: What happened?

A: Mǎrỳ níǵí-nà Pétèr.

Mary hit-NA Peter

‘MARY HIT PETER.’

Second, a significant subgroup of the West Chadic languages mark SF by subject inversion: Whenever a subject is not to be interpreted as topic, but as focus, it must occur in the prototypical focus position, i.e. in a postverbal position towards the end of the clause (e.g. Bole, (5)).

(5) Q: Who is planting the millet? A: (An) jii kàppà mòrdó yé Léngì.

3sg PROG planting millet PRT Lengi

‘LENGI is planting the millet’

4 Conclusions

The following points emerge from the discussion of (term) focus marking in Kwa, Gur and (West) Chadic:

- Focus can be marked in a variety of ways in West African languages.
- All languages in the sample exhibit a clear subject vs. non-subject asymmetry with respect to focus realizations.
- The special status of focused subjects in Kwa, Gur, and (West) Chadic follows from the fact that the subject in sentence-initial position is assigned a default topic interpretation in categorical utterances.
- If there is no prototypical match between subject and topic however, the languages of our sample show parametric variation

References

- Fiedler, I. & A. Schwarz. (2005). Out-of-focus Encoding in Gur and Kwa. In: Ishihara, Shinichiro, Michaela Schmitz and Anne Schwarz (eds.): *Interdisciplinary Studies on Information Structure 03*, 111-142. Potsdam: Potsdam University.
- Hartmann, Katharina & Malte Zimmermann (2004) Focusing in Chadic - The Case of Tangale Revisited. In: Shin Ishihara u.a. (eds) *Interdisciplinary Studies on Information Structure*. Working Papers of the SFB 632.
- Kenstowicz, Michael (1985) The Phonology and Syntax of WH-Expressions in Tangale. In *Studies in the Linguistic Sciences* 15, 79-91.
- Schwarz, Anne (2004). Aspekte der Morphosyntax und Tonologie im Buli. Mit Schwerpunkt auf dem Buli von Wiaga. Dissertation, HU Berlin.

Focus in Chadic Languages

Katharina Hartmann *Malte Zimmermann*

Humboldt University of Berlin

{k.hartmann, malte.zimmermann}@rz.hu-berlin.de

The poster gives an overview of research carried out in project B2 ‘Focus in Chadic languages’, summing up the major empirical findings and some of their implications for focus theory. The poster concentrates on the realization of focus in four West Chadic languages: Hausa, Tangale, Bole, and Guruntum.

1 Empirical findings

The empirical research into information structure and focus realization in West Chadic (Chadic, Afro-Asiatic) yields an intriguing pattern of cross-linguistic differences and similarities:

The languages under discussion do not realize focus in a uniform way, but make use of different grammatical means: syntax (Hausa), prosody (Tangale), syntax and morphology (Bole), and morphology (Guruntum).

At the same time, all languages but Guruntum exhibit a clear-cut subject/non-subject asymmetry when it comes to focus marking: subject focus must be realized, while non-subject focus can, but need not be realized. In the latter case, focus must be pragmatically resolved.

Despite their different means of realizing focus, Tangale and Bole pattern alike in that both languages realize focus on subjects by inverting the subject to a post-verbal position, similar to what has been observed for some Romance languages. Bole is special in that an optional additional morphological marker, which sets off the focus from the rest of the clause, seems to function as a

background marker rather than a focus marker, an element with similar function is found in Ewe (Ameka 1991).

Despite their different means of realizing focus, Tangale and Guruntum pattern alike in that both exhibit focus ambiguity between narrow object focus and narrow verb focus (as well as VP focus): Both foci are realized on the object NP.

2 Results and theoretical implications

The following results emerge from our investigation of focus (realization) in Chadic:

- (i.) There is no characteristic focus realization in Chadic, compared with the main accent of intonation languages. Rather, the investigated West Chadic languages realize focus in a variety of ways, notwithstanding their close genetic and areal relationship. At the same time, the cross-linguistic similarities observed suggest that the focus systems of the various languages share characteristic traits at a more abstract level.
- (ii.) The observed asymmetry between subject focus (realization obligatory) and non-subject focus (realization optional) suggests a special status for canonical, i.e. unmarked subjects as topics in these languages (Chafe 1976). It follows from their default interpretation as topics that subjects must be marked when they are not topics, but focus.
- (iii.) The optional realization of non-subject focus poses a serious challenge for standard theories of focus marking (e.g. Selkirk 1984, 1995), which make the assignment of abstract focus markers dependent on grammatical focus realization.
- (iv.) The focus ambiguity between narrow verb and object focus in Tangale and Guruntum is surprising in that it is unpredicted by current focus theories, which unanimously hold that focus is realized somewhere *on* the focus constituent and not next to it, as is the case with narrow verb focus in these languages.

More generally, our results show that a systematic research into the information structure of less well studied languages contributes to a general theory of focus in pointing out weaknesses or inadequacies of the existing accounts.

References

- Ameka, F. (1991). How Discourse Particles Mean: The Case of the Ewe “Terminal” Particles. *Journal of African Languages and Linguistics* 12. 143-170.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C. (ed.), *Subject and Topic*. London, New York: Academic Press. 25-56.
- Hartmann, K., Zimmermann, M. (2004). Focus Strategies in Chadic: The Case of Tangale Revisited. In Ishihara, S., Schmitz, M., Schwarz, A. (Eds.), *Interdisciplinary Studies on Information Structure (ISIS)*, 1. Potsdam: Universitätsverlag Potsdam. 207-244.
- Hartmann, K., Zimmermann, M. (2005a, to appear). In place-out of place? Focus in Hausa. In Schwabe, K., Winkler, S. (Eds.), *On information structure: meaning and form*. Amsterdam: Benjamins.
- Hartmann, K., Zimmermann, M. (2005b, to appear). Exhaustivity Marking in Hausa: A Re-Evaluation of the Particle *nee/cee*. In Aboh, E., Hartmann, K., Zimmermann, M. (Eds.), *Focus Strategies: Evidence from African Languages (Niger-Congo, Afro-Asiatic, West-Atlantic)*. Berlin: de Gruyter.
- Hartmann, K., Zimmermann, M. (2006, to appear). Information Structure and the Left Periphery in Dghwed’e. (Erscheint in einer Festschrift)
- Hartmann, K., Zimmermann, M. (2006, submitted). Morphological Focus Marking in Guruntum.
- Selkirk, L. (1984). *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, L. (1995). Sentence Prosody: Intonation, Stress, and Phrasing. In Goldsmith, J. (ed.), *Handbook of Phonological Theory*. Oxford: Blackwell. 550-569.

Zimmermann, M. (2006a). Adverbial Quantification and Focus in Hausa. In Ebert, C., Endriss, C. (Eds.), *Proceedings of "Sinn und Bedeutung 10"*. Berlin.

Zimmermann, M. (2006b). Focus in Western Chadic: A Unified OT-Account. In C. Davis, A.-R. Deal and Y. Zabbal (Eds.), *Proceedings of NELS 36*, UMASS, Amherst.

Zimmermann, M. (2006c). Focus Realization in West Chadic. Berlin: Humboldt Universität zu Berlin. Ms

The Role of Information Structure in Word Order Variation and Word order Change

Roland Hinterhölzl

Humboldt University of Berlin

roland.hinterhoelzl@rz.hu-berlin.de

1 Introduction

The talk focuses on word order change in Germanic. The traditional assumption is that English (and the Scandinavian languages) had a change in basic word order from OV to VO, while the continental West Germanic languages (German and Dutch) have retained the OV base order they have inherited from their Indo-European ancestors. The traditional account of this development is that the loss of Case led to a positional marking of grammatical functions in English and Mainland Scandinavian. However, this scenario is rendered implausible by the development of Dutch and Icelandic: While the former has lost its Case distinctions but retained OV order, the latter has retained its Case morphology and nevertheless changed to VO base order. A recent proposal for English assumes that the change from OV to VO is due to language contact and grammar competition (cf. Pinztuk 1999).

In this talk, I will show that the oldest stages of both English and German display OV and VO word orders, leading to the conclusion that mixed orders are not due to language contact but part of the common heritage. Thus, the question arises which factors led German and English to develop from a common base into opposite directions: While English became a pure VO language, German became a pure OV language. The second question that arises is how to account for the presence of indicators of two base orders within one language. I will dispense with the head complement parameter and show how word order

variation and word order change can be explained as a result of complex interactions in the mapping between syntactic structure and prosodic structure in which information structure plays a major role. This can be achieved in a framework that allows for competition between prosodically marked and unmarked grammatical forms.

2 Word order variation and information structure

I assume that mixed OV/VO word order is not due to grammar competition and does not signal the presence of two grammars but can be explained in terms of IS-categories. Assuming the universal base hypothesis (Kayne 1994), I will argue that word order differences cannot be accounted for by the head complement parameter or other directionality parameters and propose that the unmarked word order in a language is determined by prosodic properties, that is, the headedness of phonological phrases.

I will present two arguments showing that the differences between OV and VO languages cannot be reduced to the head complement parameter. The first observation concerns adjuncts in the middle field of English. As is illustrated in (1) and (2), adjuncts that can occur between the subject and the VP in English are subject to conditions that are absent in OV-languages, namely, the requirement that the head may not contain material to its right.

The second observation concerns the distribution of event-related adjuncts. As is illustrated in (3), Time, Place and Manner adjuncts occur preverbally in the order T>P>M in OV languages but postverbally in the inverted order in VO-languages. I have argued in Hinterhölzl (2004) that the preverbal order of these adjuncts corresponds to their base order and that the inverted order in English is due to successive cyclic VP-intrapolation that was brought about by a

stylistic rule of light verb raising that primarily applied to event related adjuncts in OE, since they are typically realized as (rather heavy) NPs and PPs.

Then I will argue that the same prosodic restriction that applies in the I-domain in English not only occurs in VO-languages but is a universal constraint that also applies in OV-languages and has been restricted to the V-domain in the history of German. Assuming that heaviness can be defined in terms of prosodic branching the pertinent condition can be formulated as a constraint on the mapping between syntactic structures and prosodic structures, as is given in (4). Since this condition, as a typical interface condition, is violable and yields ungrammatical results only if no alternative structure is available, it serves well in explaining word order tendencies in historical developments.

3 Information structure and word order change

The final part of the talk addresses the question of how German became a pure OV language. First I will show that word order in OHG was determined by IS-conditions, as is illustrated in (5), with contrastive foci being primarily realized in preverbal position. The positioning of focussed constituents is important for determining the unmarked word order in the clause, since due to focus restructuring preverbal narrow focus strengthens the prosodic pattern (s w), while postverbal narrow focus strengthens the prosodic pattern (w s). Then I will argue that the change in unmarked word order in German was brought about by the grammaticalization of the definite determiner that was first introduced with discourse given NPs in the preverbal domain. When the use of the definite determiner was extended to stressed discourse new NPs just in case they were uniquely identifiable via bridging or binding or due to their semantic properties, their preverbal positioning (according to the initial pattern)

strengthened the prosodic pattern (s w) and introduced right-branching phrases in the middle field such that over time the mapping condition in (4) was restricted to the V-domain in German.

- (1) a. John (more) often (* than Peter) read the book
 b. Hans hat öfter (als der Peter) das Buch gelesen
- (2) a. John read the book more often than Peter
 b. * Hans hat das Buch gelesen öfter (als Peter)
- (3) a. C T P M-V OV-languages
 b. C V- M P T VO-languages
- (4) A right-headed phonological phrase must sit on a right branch with respect to the syntactic head that is to become its prosodic sister
- (5) C background V focus

References

- Hinterhölzl, Roland (2004). "Language Change versus Grammar Change". In Eric Fuss & Carola Trips (eds.): *Diachronic Clues to Synchronic Grammar*. Linguistik Aktuell 72. Amsterdam.
- Pintzuk, Susan (1999). *Phrase Structures in Competition: variation and change in Old English word order*. New York: Garland. John Benjamins.

Focus in Gur and Kwa

Ines Fiedler, Brigitte Reineke, Anne Schwarz

Humboldt University of Berlin

sfb632.b1@rz.hu-berlin.de

1 Investigation and Language Sample

The project investigates into focus phenomena in the two genetically related West African Gur and Kwa language groups of the Niger-Congo phylum. They are similar with respect to word order typology (all are SVO languages), but they are of divergent morphological type (agglutinating Gur versus isolating Kwa), and most are tone languages.

A major contribution of our project within the SFB 632 consists of supplying information and analysis on focus expressions from the perspective of typologically divergent non-European tone languages. Our studies concern among others the following fields of tasks:

- Range of interferences between grammatical structure and focus strategies on the basis of language typological comparison
- Analysis of focus marking from a diachronic perspective

Apart from studying the available literature and data contributions by colleagues working on related languages, the following Gur and Kwa languages have been examined so far by the project members:

- Gur: Buli, Byali, Dagbani, Ditammari, Gurene, Konkomba, Konni, Nateni, Yom
- Kwa: Akan, Efutu, Foodo, Gbe (Ewe, Aja, Fon), Lelemi

The poster displays some selected results that have come out of our studies so far.

2 Typological influence on focus expressions

2.1 Phonological type

Preliminary checks in some of the sample languages (Ewe, Fon, Buli, Dagbani, Konni) suggest that different from Indo-European non-tonal languages, fundamental frequency / pitch does not function as significant focus indicator.

2.2 Morphological type

While all languages considered provide syntactically unmarked (i.e. *in-situ*) focus strategies for verb and / or postverbal complements, there is a difference whether morphological marking is involved. The sample languages vary according to their morphological type: agglutinating Gur languages often use morphological focus markers (underlined), isolating Kwa languages don't:

- (1) (a) What did the woman eat? – She ate BEANS
 (b) What did the woman do? – She ATE BEANS.

KWA	GUR
<i>Efutu</i> mú-dì èdùbá.	<i>Byali</i> ù hwábé tūnā è.
<i>Ewe</i> é ɖù àyī.	<i>Ditammari</i> ò dī yātūrà <u>n̄yā</u> .
<i>Fon</i> é ɖù àyìkún.	<i>Konkomba</i> ù ŋmá ŋētũ̀n̄ <u>lā</u> .
<i>Lelemi</i> ó-dì àkábí.	<i>Dagbani</i> ò ŋùbì- <u>lá</u> tùyà.
<i>Foodo</i> òó wí ácéè.	<i>Buli</i> ò ŋòb <u>kà</u> túé.
<i>Akan</i> ò-dì àdùá.	<i>Konni</i> ù ŋòbì- <u>wá</u> túò.

3 Extra-clausal focus constructions

Apart from the well-known fact that cleft constructions are used to express focus (e.g. Byali, cf. Reineke), we observed striking similarities between the out-of-focus part of *ex-situ* term focus constructions and narrative clauses in many languages irrespective of the language specific structural properties (clause-initial conjunction, special pronominal forms, verb suffixes, tone, etc.; cf. Fiedler & Schwarz). Akan for example, displays the same conjunction *nà* and tonal changes at the predicate.

(2) Akan

(a) focus construction:

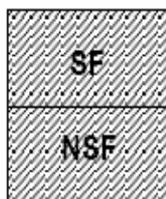
àdùá nà ò-díì-yé.
 beans CNJ 3sg-eat-YE
 ‘He ate BEANS.’

(b) narrative construction:

máàmé nòáà àdùá, nà n-àdámfùò díì-yé.
 Maame cook beans CNJ POSS:3sg-friend eat-YE
 ‘Maame cooked beans and her friend ate them.’

We conclude that this recurrent parallelism is due to a systematic extra-clausal construction containing a narrative clause following the focus constituent (narrative hypothesis), even though the distribution of the narrative structures within *ex-situ* focus constructions differs within our language sample: in many Kwa languages, the narrative pattern is found in non-subject as well as subject focus constructions. In many Gur languages, subject focus constructions do not participate in the narrative pattern.

Akan, Ewe, ...



Buli, Dagbani, ...



narrative pattern

SF subject focus construction

NSF *ex-situ* non-subject term
focus construction

References

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

<http://www.sfb632.uni-potsdam.de>

Reineke, Brigitte. to appear. Identificational Operation as Focus Strategy in Byali. In: Aboh, Enoch, Katharina Hartmann & Malte Zimmermann (eds.), *Focus Strategies: Evidence from African Languages*, Berlin: de Gruyter.

Discourse Structure and Information Packaging in Cross-Linguistic Perspective

Svetlana Petrova Anne Schwarz

Humboldt University of Berlin

s.petrova@staff.hu-berlin.de anne.schwarz@rz.hu-berlin.de

1 Linguistic Relevance of Discourse Organization

Part of the investigation within projects B1 and B4 is concerned with the interaction between Information Structure and Discourse Semantics for the explanation of different phenomena in genetically non-related languages. We observed that similar principles of discourse organization can be traced as relevant for structural variation in early Germanic languages as well as in West African languages of the Gur group. In early Germanic, distinctions between sentences on purely discourse-related considerations are responsible for the placement of the inflected verb while in some languages of the Gur group, this is reflected in verb morphology as well as in the selection of connectives in context.

2 Discourse Basis for Verb Placement in Early Germanic (B4)

Results of earlier studies (Hinterhölzl et al. 2005) indicate that the inflected verb in Early Germanic enters the borderline between **old and new information** in an utterance and thus **marks the beginning of the domain of new-information focus**. This phenomenon is best shown on the complementary distribution of V/1 vs. V/2 in sentences of the presentational vs. categorical kind. A problem apparently arises with sentences with **discourse-given referents** but nevertheless exhibiting V/1 as in (1):

- (1) Lat. Pharisaeus autem **coepit** intra se/ reputans dicere
 OHG **bigonda** ther phariseus innan imo/ ahtonti quedan (T 126, 5–6)
 ‘The Pharisee started to speak thinking by himself ‘

On closer inspection it turns out that patterns like the one in (1) regularly occur under the following conditions: i) **positional restrictions**: in discourse-initial sentences or in episode-onsets; ii) **with special classes of verbs**, most commonly verbs of motion and saying as well as inchoative verbs, especially perception verbs. Thus, we conclude that such instances signal the initiation of a new episode or sub-episode in the structure of a running text or a shift in the deictic orientation of the narrative setting (place, time, or participant). The use of V/1 in such sentences – quite similar to the one in presentational ones – fulfils the following functions:

- it disables a topic-comment structure: the given referent should not receive a topic interpretation;
- it signals broad/ sentence focus.

Following these observations on the distribution of V/1 vs. V/2 in OHG, we conclude that verb syntax is ruled by discourse factors in Early Germanic. Implementing the model of discourse semantics proposed by Asher & Lascarides (2003) we are able to define the following functional distribution of verb patterns in early Germanic:

- V/1 indicates relations of coordination (Narration) or the shift from a lower level of subordination / continuation to a higher level of discourse hierarchy, ex. the end of a continuation chain and the return to the level of main story;

- V/2 indicates relations of subordination and continuation, i.e. it develops deeper, supportive parts of discourse, ex. Elaboration and Explanation on a discourse given participant in categorical topic-comment sentences.

Comparative work reveals that the contexts triggering regular V/1 placement in OHG are the same in the other early Germanic languages as well. Nevertheless, there are differences concerning the realization of the utterances belonging to the subordinating type of linking which in turn can be made responsible for the grammaticalization of different word order patterns in English vs. the continental West-Germanic group.

3 Discourse Basis for Morphosyntactic Variation in Gur (B1)

Recent research points out that some languages of the Gur group (Buli, Konni, Dagbani, and others) use morphosyntactic means (verb morphology, connectives etc.) to express a discourse based difference between clauses with or without a topical subject: In the canonical sentence construction (SVO, 2a) the subject represents the sentence topic about which a comment is made within the clause. Any deviation from this categorical configuration occurs in morphosyntactically marked constructions in which the predicate is either hypotactically (2b) or paratactically (2c) encoded.

(2) Topic in Sentence and Discourse

- [+ topical subject]: a. **canonical predicate: SV(O)**
 categorical utterance with topic + comment
- [- topical subject]: b. **hypotactic predicate (incorporation): SV(O)**
 thetic utterance without topic
*descriptive backgrounding, major protagonists,
 affinity with imperfective aspect*
- c. **paratactic predicate (coordination): OSV**
 topic + clausal comment
*sequential foregrounding, narrative story line,
 affinity with perfective aspect*

The discourse based grounding function of the two marked configurations (2b/2c) becomes transparent from their distribution in (narrative) discourse:

- The hypotactic predicate is typically used text-initially, where major discourse referents are being introduced by verbs of existence or posture in order to provide background for the upcoming string of plot. Accordingly, we find the same thetic construction in other backgrounding contexts, like temporal and relative clauses, but also with subject and sentence focus.
- The paratactic predicate on the other hand is typically found with important events making up the major story line in the foreground. In this complex sentence construction involving a clause boundary between initial constituent and following subject, the sentence-initial constituent rather than the subject of the following clause serves as unexpected new or contrastive topic.

Interesting in the Gur sample are interactions of these constructions with other grounding devices, like aspect (cf. Hopper 1979). The thetic construction (2b) allows the speaker to present major (animate) discourse referents

irrespective of any remarkable event in the story line (individual-level predicates) and shows special imperfective features; while with the other marked construction (2c) the speaker denotes a transitional relation between the referent of the sentence-initial constituent and the predicate of the following clause (stage-level predicates), which fades as soon as the plot of the story develops further.

References

- Asher, Nicholas and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.
- Fiedler, Ines and Anne Schwarz. 2005. Out-of-focus Encoding in Gur and Kwa. In Shinichiro Ishihara, Michaela Schmitz and Anne Schwarz (eds) *Interdisciplinary Studies on Information Structure 03*, Potsdam: Potsdam University, 111-142.
- Hinterhölzl, Roland and Svetlana Petrova. 2005. Rhetorical Relations and Verb Placement in the Early Germanic Languages. Evidence from the Old High German Tatian translation (9th century) In Manfred Stede et al. (eds) *Salience in Discourse. Multidisciplinary Approaches to Discourse*, Münster: Stichting/Nodus, 71–79.
- Hopper, Paul J. 1979. Aspect and Foregrounding in Discourse. In Talmy Givón (ed) *Syntax and Semantics*, New York: Academic Press, 213-241.

Information Structure and Word Order in the Early Germanic Languages and its Analysis in a Linguistic Database

Karin Donhauser Roland Hinterhölzl Svetlana Petrova, Michael Solf
Humboldt University of Berlin

Manfred Stede Stefanie Dipper Michael Götze Julia Ritz
University of Potsdam

karin.donhauser@rz.hu-berlin.de

1 Introduction

Project B4 of SFB 632 investigates the interaction between information-structural and grammatical conditions in the development of word order regularities in the Germanic languages. The central issue of research is how the information-structural role of a constituent influences its positioning in the clause. The results of this study shed light on the complex interaction between information structure and core grammar and enable us to make important contributions to the theory of language change. Together with project D1 (“Linguistic Database for Information Structure”), we employ corpus-linguistic methods both for creating, representing and analyzing data of early Germanic languages. In this poster, we present major aspects of our corpus work, focusing on the process of data annotation and its analysis in a linguistic database.

2 Annotation and technical infrastructure

We annotate morpho-syntactic and information-structural features in texts from the Early Germanic period. For the purpose of annotating information-structural phenomena on texts available in written form only, we developed an annotation

scheme based on a multi-layer model of information-structural representation as proposed in Molnár (1993):

- Pragmatic status of discourse referents: given, new, accessible
- Predicational structure: Topic vs. Comment
- Informational relevance: Focus vs. Background

For digitization and annotation, we use EXMARaLDA⁴, a user-friendly annotation tool for discourse annotation with an underlying XML-format. The resulting data is converted into a generic standoff representation format PAULA (“Potsdamer Austauschformat für Linguistische Annotationen“, cf. Dipper 2005), which allows for the representation of both hierarchical and time-aligned annotations. For exploration and data analysis, “ANNIS: a linguistic database for Annotated Information Structure”⁵ – a web application accessible through standard web browsers for visualizing and querying the data – is employed.

3 Interactions between information structure and verb placement

We demonstrate our method on a case study completed on the Old High German Tatian translation. The corpus comprises sentences displaying a different word order in relation to the Latin original. We look for the correlation between verb placement and informational status of discourse referents in cases of the presentational/thetic vs. categorical distinction (Sasse 1995), cf. (1)–(2) vs. (3):

⁴ <http://www1.uni-hamburg.de/exmaralda>

⁵ <http://www.sfb632.uni-potsdam.de/annis>

Q1: Query for the introduction of new discourse referents

context=*Ersteinführung* & clause-status=main*
query for first mention (Ersteinführung) in main clauses

Q2: Query for 4 i): "FOC[Vfin ... DRnew ...]" in the results of Q1

(context=*Ersteinführung* & clause-status=main*) AND
 ((!pos=adv* & clause-status) OR (pos=adv* & cat=vp & !#5 .* #6))
either there is no adverb in the clause OR the adverb does not precede the vp

Q3: Query for 4 ii): "Frame FOC[Vfin ... DRnew ...]" in the results of Q1

(context=*Ersteinführung* & clause-status=main*) AND
 (pos=adv* & cat=vp & #3 .* #4)
an adverb precedes the vp

In the Tatian corpus (1748 sentences), both patterns occur nearly equally often (Q1: 12 hits, Q2: 6 hits; Q3: 5 hits; one sentence contained an annotation error). This provides evidence for variation within the functional domain of presentational sentences, which indicates that we encounter conditions for a language change situation in this functional domain (cf. Lightfoot 1999).

5 Future Work

Future work includes the optimization of current annotation and the development of more scenarios for queries, and statistical validation of several working hypotheses. One of them concerns the role of discourse-level phenomena for the realization of different syntactic patterns in the early Germanic dialects as first observed for OHG in Hinterhölzl & Petrova (2005). Following this, we plan to provide additional annotation of discourse structure on texts from different early Germanic dialects using specialized annotation tools (ex. RST-tool).

References

Dipper, Stefanie. 2005. XML-based Stand-off Representation and Exploitation of Multi-Level Linguistic Annotation. In *Proceedings of Berliner XML Tage 2005 (BXML 2005)*, 39-50, Berlin, Germany.

-
- Hinterhölzl, Roland and Svetlana Petrova. 2005. Rhetorical Relations and Verb Placement in the Early Germanic Languages. Evidence from the Old High German Tatian translation (9th century) In Manfred Stede et al. (eds) *Salience in Discourse. Multidisciplinary Approaches to Discourse*, Münster: Stichting/Nodus, 71–79.
- Lightfoot, David. 1999. *The Development of Language. Acquisition, Change, and Evolution*. Blackwell.
- Molnár, Valería. 1993. Zur Pragmatik und Grammatik des TOPIK-Begriffes. In *Wortstellung und Informationsstruktur*, hg. M. Reis, 155–201. Tübingen: Niemeyer.
- Sasse, Hans-Jürgen. 1995. “Theticity” and VS order: a case study. *Sprachtypologie und Universalienforschung* 48, 1/2: 3–31

Information Structure and the Anticipation of Discourse Referents

Thomas Weskott Robin Hörnig* Elsi Kaiser⁺ Caroline Féry**
Gisbert Fanselow Sabine Kern* Reinhold Kliegl**

*University of Potsdam

University of Southern California, Los Angeles

weskott@uni-potsdam.de

Different languages use different means—including articles, word order and intonation—to encode the distinction between GIVEN and NEW information. Kaiser & Trueswell (2004) investigated one of these information-structural options, word order, and found that in Finnish, noncanonical OVS sentences trigger the anticipation that the referent of the upcoming postverbal noun will be NEW rather than GIVEN.

We investigate whether an arguably more subtle type of information-structural marking, namely intonation, also triggers anticipatory effects. Intonation differs from word order insofar as the propositional meaning of a sentence can be understood even if intonation is ignored, whereas ignoring overt marking of word order (case-marking on the determiner) can cause misinterpretations, especially with noncanonical orders. To determine whether this difference has implications for processing, we examine whether K&T's word order findings replicate for German (Exp.1), and whether intonation, which in German has been claimed to provide cues to discourse-status, triggers comparable anticipatory effects (Exp.2). In a third experiment, we compared the anticipatory effects triggered by intonation and by word order.

In all three experiments, while listening to auditory stimuli (examples (1), (2)), participants saw scenes depicting three referents, e.g., doctor, nurse, patient. Two referents were mentioned prior to the target sentence (i.e., GIVEN).

The third was discourse-NEW. We tracked the time-course of participants' attention to referents in the scenes.

Exp.1, Word order (SVO/OVS). We followed K&T in hypothesizing that, if OVS in German serves to signal the discourse-newness of the postverbal referent, OVS sentences (example (1.b)) will trigger more anticipatory fixations to NEW referents than canonical SVO sentences (ex.1a). Crucially, this is predicted to happen before the postverbal referent is mentioned.

Exp.2, Intonation. We tested SVO sentences with two intonational patterns (examples (2a-b)). We predicted that an early falling tone (example (2a)) will trigger anticipatory fixations on GIVEN referents, because in German, after a falling tone only GIVEN material can occur postverbally. However, an early rising tone (example (2b)) should trigger anticipatory fixations on NEW referents. In these kinds of sentences in German, a second accent, with a falling pattern, is expected later in the sentence. All sentences were declaratives, which in German have a global falling contour.

We also manipulated the discourse-status (GIVEN/NEW) of postverbal referents in both experiments. We predicted that if there is a mismatch between the discourse-status of the mentioned referent and that predicted by intonation/word order, participants would have to reallocate attention, which would delay fixations to the mentioned referent.

The predictions were borne out. We found an early effect of word order in Exp.1, and an early effect of intonation in Exp.2. OVS word order (example (1b)) triggered the expectation that the upcoming referent is new, as did the intonation contour with the late fall (example (2b)). The magnitudes of the

effects were strikingly similar in the two experiments, although the effect of intonation appeared somewhat later and weaker. Effects of the discourse status of the actually mentioned entity appeared at a later point in time. These results are in accordance with sentence-completion experiments that we have conducted.

Exp.3, Comparison of Word Order and Intonation. In this experiment, we tested how the anticipatory effects triggered by intonation fare in the presence of the word order manipulation. To this end, we compared the intonational variants early fall/new referent mentioned and late fall/new referent mentioned with the word order conditions SVO/given referent mentioned and OVS/given referent mentioned. Based on the results of Exps. 1+2, we hypothesized that the reallocation effect triggered by intonation will be present despite the presence of the word order manipulation.

These predictions were borne out. Although we observed an effect of intonation at an early point, the effect of word order was stronger and more persistent across time. We conclude that intonational effects survive the presence of the word order manipulation, but that the latter exert a stronger influence on participants' expectations about the discourse status of referents. Taken together, our results suggest that two different means of information-structural marking—word order and intonation—have remarkably similar effects on the time course of the expectations that listeners build up concerning the discourse-status of upcoming referents.

Materials

Sample Context:

An der Empfangstheke eines Krankenhauses lehnen ein Arzt und eine Krankenschwester. // Die Uhr zeigt fast zwei.

A doctor and a nurse are leaning on the reception desk of the hospital. // It's almost two o'clock.

(1) *Target Sentence, Exp. 1:*

(a) SVO: Der Arzt befragt gleich die Krankenschwester_{GIVEN}/die Patientin_{NEW}.
The-NOM doctor will soon question the nurse/the patient.

(b) OVS: Den Arzt befragt gleich die Krankenschwester_{GIVEN}/ die Patientin_{NEW}.
The-ACC doctor will soon question the nurse/the patient.

(2) *Target Sentence, Exp. 2::*

(a) early fall: Der Arzt be\FRAGT gleich die Krankenschwester_{GIVEN}/die Patientin_{NEW}.

(b) late fall: Der Arzt befragt gleich die \KRANKenschwester_{GIVEN}/die Pa\TIENtin_{NEW}.

References

Kaiser, E. & Trueswell, J.C. (2005). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2), 113-147

Contextual Licensing of Marked Word Order in Spatial Descriptions

*Robin Hörnig Thomas Weskott
Reinhold Kliegl Gisbert Fanselow*

University of Potsdam

rhoernig@uni-potsdam.de

According to mental model theory (Johnson-Laird, 1983), when reading a spatial relational sentence like *The deer is to the left of the snake*, the reader constructs an initial mental model of the described layout. When the context sentence is followed by another such sentence, the target sentence, e.g., *The donkey is to the right of the snake*, the reader will try to integrate the new entity introduced by the target sentence, *the donkey*, into the initial model. Hörnig, Oberauer, and Weidenfeld (2005) propose the relatum = given principle, according to which model integration is easier if the relatum of the target sentence (the prepositional object *the snake*) is already given within the initial model, while the locatum of the target sentence (the grammatical subject *the donkey*) is the new entity. Relatum = given, together with the given-new principle (cf. Clark & Haviland, 1977), predicts that model integration will be easiest with the given relatum in preverbal position, although the word order of such a sentence is marked (e.g., *To the right of the snake is the donkey*). Hörnig et al. tested this prediction by manipulating givenness (relatum vs. locatum of the target sentence) and word order (unmarked vs. marked) in German. The two sentences of a pair, which were always of the same word order, were presented one after another self-paced by participants. Comprehension times for the target sentence were taken as indicating ease of model integration. Integration was easiest with a given preverbal relatum and most difficult with a new preverbal

relatum, with both sentences bearing a marked word order. With unmarked word order, integration was of intermediate difficulty. This finding might be interpreted in terms of given-new licensing of marked word order, i.e., sentences with marked word order are easy to comprehend, as long as their order is given-new.

We offer and test an alternative interpretation in terms of poset (= partially ordered set) licensing inspired by Prince (1999). On the poset account, marked word order is licensed if the preverbal constituent denotes a new entity which is poset-related to a given entity. According to our analysis, the preverbal prepositional phrase as a whole (*to the right of the snake*) denotes a discourse-new place that is spatially related (i.e., poset-related) to the given place occupied by the snake. On the poset account, in order to be licensed, the spatial preposition has necessarily to be part of the preverbal constituent. On the given-new account, it is sufficient that the given preverbal relatum precedes the new locatum. In order to evaluate the two accounts, we reran the experiment of Hörnig et al. with spatial adverbs instead of spatial prepositions. In particular, we realized four different word order variants with spatial adverbs and crossed them with givenness (relatum vs. locatum). Examples of the four word order variants with a given relatum are presented below (givenness of the locatum is achieved by mentioning the donkey instead of the snake in the context sentence, e.g., *The donkey is as from the deer left* for (1)):

- (1) *Der Hirsch ist von der Schlange aus links. Der Esel ist von der Schlange aus rechts.*
'The deer is as from the snake left. The donkey is as from the snake.'
- (2) *Von der Schlange aus ist der Hirsch links. Von der Schlange aus ist der Esel rechts.*
'As from the snake the deer is left. As from the snake the donkey is right.'
- (3) *Von der Schlange aus links ist der Hirsch. Von der Schlange aus rechts ist der Esel.*
'As from the snake left is the deer. As from the snake right is the donkey.'
- (4) *Links von der Schlange aus ist der Hirsch. Rechts von der Schlange aus ist der Esel.*
'Left as from the snake is the deer. Right as from the snake is the donkey.'

Are all of the target sentences with marked word order in (2) to (4) contextually licensed due to the given relatum? The critical sentence is (2), with the relatum being preverbal as in (3) and (4), but with the spatial adverb being sentence final as in the sentence with unmarked word order in (1). On the given-new account, (2) is as licensed as are (3) and (4). On the poset account, (3) and (4) are licensed but (2) is not. Comprehension times for the target sentence revealed that model integration is easier with (2) as compared to (1), as predicted by given-new. However, in accordance with poset, model integration is further facilitated with (3) and (4) as compared to (2). We conclude that contextual licensing of marked word order is a matter of degree. Given-new was sufficient to licence marked word order. Poset, however, licensed marked word order to a larger extent than given-new did. In terms of model integration, comprehension seems to further benefit from the early availability of relational information which specifies where to add the new place to the model. Taken together, our results suggest that linguistic proposals on licensing of marked word order might well meet with requirements identified for cognitive processing, e.g., mental integration.

References

- Clark, H. H., & Haviland, S. (1977). Comprehension and the given-new contract. In R. O. Freedle (ed.), *Discourse processes: advances in research and theory, Vol. 1: discourse production and comprehension* (pp. 1-40). Norwood, NJ: Ablex.
- Hörnig, R., Oberauer, K., & Weidenfeld, A. (2005). Two principles of premise integration in spatial reasoning. *Memory & Cognition*, 33, 131-139.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA : Harvard University Press.

Prince, E. (1999). How not to mark topics: 'topicalization' in English and Yiddish. *Texas Linguistics Forum*. Austin: University of Texas.

Structural Markedness and Structural Facilitation in Comprehending Descriptions of Unknown and Known Spatial Layouts

Robin Hörnig, Thomas Weskott, Reinhold Kliegl, and Gisbert Fanselow
University of Potsdam

rhoernig@uni-potsdam.de

In two eye-movement experiments we investigated contextual licensing of marked word order in German spatial relational target sentences and examined to what extent contextual licensing benefits from a parallel word order of a preceding context sentence. Previous research has shown that a spatial relational target sentence with given-new order – as established by a preceding context sentence – is easier to comprehend when word order is marked, as in (1b), rather than unmarked, as in (1a) (cf. Hörnig, Oberauer, & Weidenfeld, 2005).

- (1) a. *Es gibt einen Hirsch. Der Hirsch ist über der Schlange. Sie ist über dem Esel.*
,There is a deer. The_{masc} deer is above the_{fem} snake. It_{fem} is above the donkey.’
b. *Es gibt einen Hirsch. Unter dem Hirsch ist die Schlange. Unter ihr ist der Esel.*
,There is a deer. Below the_{masc} deer is the_{fem} snake. Below it_{fem} is the donkey.’

We explain the contextual licensing of a marked target sentence in terms of poset licensing (poset: partially ordered set): Marked word order of a target sentence with a preverbal given relatum as in (1b) is licensed because the preverbal prepositional phrase as a whole denotes a new place occupied by a new entity (e.g., the donkey) that is poset-related (here: spatially related) to a given place occupied by a given entity (e.g., the snake) (cf. our other poster in this session). Poset licensing, which is stronger than mere given-new licensing, corresponds on the processing side to an easier integration of a sentence final new locatum (the donkey in (1b)) rather than a sentence final new relatum (the

donkey in (1a)). Our two experiments differed in the supposed strength of poset licensing. In the first experiment, participants read sentence pairs describing unknown layouts, whereas in the second experiment, participants saw a depiction of a layout before they read the sentence pair describing the layout, which was then known. The new entity mentioned last in the target sentence was accordingly either completely new (unknown layouts) or known but not previously mentioned (known layouts). Since mentally integrating the new element should be necessary only for unknown layouts but not for known layouts, we expected poset licensing to be stronger in Experiment 1 than in Experiment 2. In particular, integration in Experiment 1 should be reflected in gaze durations on the sentence final noun phrase denoting the new entity. The context sentences in our experiments had a given-new order, too. The given-new order of a context sentence was established by a preceding introductory sentence of the form *Es gibt ein x* ‘There is an x’ with the introduced entity x being always denoted by the first noun phrase of the context sentence. Unlike the context sentences themselves, an introductory sentence simply mentioned an entity x without localizing it. Since the place of x was known in Experiment 2 but not in Experiment 1, the licensing strength for marked context sentences in the two experiments should be opposite to the licensing strength for marked target sentences. Marked context sentences describing a known layout, which are poset licensed, should be licensed more strongly than context sentences describing an unknown layout, which are given-new licensed.

In the Hörnig et al. (2005) study, all sentence pairs were of parallel word order. So we cannot assess to what extent the strong licensing effect has to be attributed to structural parallelism. In a production experiment, Hartsuiker, Kolk, and Huiskamp (1999) have attested a structural priming effect for Dutch spatial relational assertions. On the other hand there is evidence from a

comprehension study by Frazier, Taft, Roeper, Clifton, & Ehrlich (1984) that structural facilitation induced by parallel structures in conjoined sentences more probably affects marked than unmarked constructions. In addition to sentence pairs with parallel word order we also tested sentence pairs with nonparallel word order as those in (2a) and (2b). We expected to obtain a structural facilitation effect for target sentences when preceded by a context sentence with parallel word order, an effect that might be confined to marked word order.

- (2) a. *Es gibt einen Hirsch. Unter dem Hirsch ist die Schlange. Sie ist über dem Esel.*
 ‚There is a deer. The_{masc} deer is above the_{fem} snake. It_{fem} is above the donkey.’
 b. *Es gibt einen Hirsch. Der Hirsch ist über der Schlange. Unter ihr ist der Esel.*
 ‚There is a deer. Below the_{masc} deer is the_{fem} snake. Below it_{fem} is the donkey.’

In both experiments we found a substantial structural facilitation effect in both measures, self-paced total inspection times for sentence pairs and gaze durations on target sentences. Structural facilitation turned out to be strictly confined to target sentences which were of marked word order. As predicted, contextual licensing for marked target sentences was stronger for unknown than for known layouts. The processing advantage for marked target sentences showed up in gaze durations on the sentence final noun phrase but not earlier in the sentence. As regards marked context sentences, contextual licensing was stronger for known than for unknown layouts. Beyond the evidence that parallel sentence pairs help in processing marked word order there was also evidence that nonparallel sentence pairs hurt in processing marked word order. In Experiment 2, gaze durations were shorter on marked than on unmarked context sentences. However, gaze durations were in no instance shorter on marked than on unmarked target sentences unless they did benefit from a parallel word order of the preceding context sentence. We conclude that an unmarked context sentence hampered the processing of a following marked target sentence. Our

findings show that both contextual licensing as well as structural parallelism affect the processing of marked word order to a considerable degree.

References

- Frazier, L., Taft, L., Roeper, T., Clifton, C., & Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition*, 12, 421-430
- Hartsuiker, R. J., Kolk, H. H. J., & Huiskamp, P. (1999). Priming word order in sentence production. *The Quarterly Journal of Experimental Psychology*, 52A, 129-147.
- Hörnig, R., Oberauer, K., & Weidenfeld, A. (2005). Two principles of premise integration in spatial reasoning. *Memory & Cognition*, 33, 131-139.

The Acquisition of the Additive Focus Particle *auch* in German: A Matter of Competence or Performance?

*Frauke Berger*¹ *Anja Müller*² *Barbara Höhle*¹
*Michaela Schmitz*¹ *Jürgen Weissenborn*²

¹ University of Potsdam ² Humboldt-University Berlin

1 Introduction

In the present study we investigate the abilities of German children to produce and to process sentences containing the focus particle *auch* ('also'). The additive focus particle *auch* signals that one of the constituents of the sentence it appears in is added a set of alternatives given in the discourse model. The identification of the constituent that is to be interpreted in relation to a set of alternatives (therefore the particle's domain of application) depends on positional and prosodic information.

If *auch* is located in postverbal position of a sentence in SVO-order, the particle potentially can apply to 2 different constituents (at least):

- 1a) *[Toby]_{DoA} hat AUCH eine Puppe.* ' *[Toby]_{DoA} has ALSO a doll.*'
 'There are other people than Toby possessing a doll'
- 1b) *Toby hat auch [eine PUPPE]_{DoA}.* 'Toby has also [**a DOLL**]_{DoA}.'
 ⇒ 'Toby possesses a doll besides other things'

In 1a) the preceding subject (*Toby*) is the domain of application and the particle itself is accented. In contrast, in 1b) the following object (*eine Puppe*) is the domain of application. In this case the object is accented and the particle is unaccented.

Previous acquisition studies found that children up to school age still have difficulties comprehending sentences containing the focus particle *auch* (Hüttner

et al. 2004; Bergsma 2006 for *ook* (D)). Following Hüttner et al., sentences with unaccented *auch* in postverbal position are often interpreted as sentences with accented AUCH, suggesting a default interpretation in favour of the latter one. This late receptive mastery of *auch* stands in contrast to its early production already in children's first multi-word utterances (Nederstigt 2001; Penner et al. 2000) suggesting that the acquisition of the focus particle could be an instance of a pattern of productive knowledge preceding interpretative knowledge, which is unusual.

2 Experiments

To further test this suggestion we looked more closely both on the comprehension and the production of sentences containing *auch* by German learning children. In a production-experiment we addressed the question whether 28- and 32-month-old children are able to compute the focus structure of a sentence in a given context and would produce the accented and the unaccented *auch* adequately. We tested 14 children at the age of 28 months. 12 of them were tested again at the age of 32 months. In order to elicit the production of the target structures as “*auch /AUCH einen Ball*” (‘*also /ALSO a ball*’) we used a task in which a set of pictures had to be described by a sentence completion task.

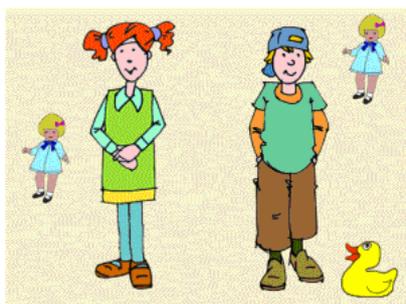
The data analysis showed that the 28-month-old children did not produce any unaccented *auch*-utterances, but in contrast we found in 27% of the cases the accented AUCH+NP. At an age of 32 months we could elicit the unaccented *auch*+NP in 19% of all cases and the accented AUCH+NP in 73% of all cases.

These results suggest an advantage for the production of utterances containing the accented *auch* reflected by an earlier appearance already for the

28-month-olds and by the still more frequent use by the 32-month-olds. Nevertheless we would argue, that even the 28-month-olds can differentiate between contexts which require the use of an accented *auch* and those that require an unaccented *auch*. This is based on the observation that the children produced the accented AUCH just in the proper contexts but never in a context in which an unaccented *auch* is required. We will discuss our observations in relation to the results of previous production studies and point out potential linguistic and extra-linguistic factors that might be responsible for the pattern of results.

The results of the 2nd study that we present are questioning the claim of previous comprehension studies that children up to school age have difficulties in interpreting sentences with the focus particle *auch*. In contrast to off-line methods (such as sentence-picture-matching tasks and felicity-judgment tasks) used in previous experiments we adopted an eye-tracking paradigm that allows us to track the processing of the input without the need of any conscious reaction by the child.

34 4-year-olds were presented with SVO sentences either containing the accented postverbal *auch*, the unaccented postverbal particle *auch* or no particle at all. As visual stimuli, two children were displayed on a monitor (see fig.1). One child had one object (e.g. a doll), the other child had two objects, one of them being identical to the other child's (e.g. a doll and a duck). The child with the two objects served as subject in all sentences (cf. sentence 1a) and 1b)).



We hypothesized that the processing of the different kinds of information structure in test sentences should result in different fixation preferences of the corresponding alternative sets given in the visual scene.

The experimental results are consistent with our predictions. Children directed significantly longer fixations to the alternative set for the subject of sentences with the accented AUCH, as compared to sentences with an unaccented *auch* or with no *auch* at all. Similarly, children fixated the alternative set for the object of the sentence more often with sentences containing unaccented *auch* as compared to accented AUCH.

The comprehension results indicate that children as young as 4 years display more adult-like comprehension of sentences with the focus particle *auch* than indicated by previous studies. The children in our study show the ability to establish alternative sets of sentence constituents in the discourse model (cf. Paterson et al. for *only*). Furthermore, they seem to be able to use accent as a cue for determining the appropriate domain of application of the postverbal focus particle *auch*. This suggests that children's performance in previously used off-line techniques/tasks might underestimate children's abilities/competence in this area.

References

- Bergsma, W. (2006): (Un)Stressed *ook* in Child Dutch. In V. van Geenhoven (ed.) *Semantics in Acquisition*. Dordrecht: Springer, 329-348.
- Hüttner, T.; Drenhaus, H.; van de Vijver, R. & Weissenborn, J. (2004). The Acquisition of the German Focus Particle *auch* 'too': Comprehension does not Always Precede Production. In *BUCLD 28: Proceedings of the 28th annual Boston University Conference on Language Development*. On-line Supplement.
- Nederstigt, U. (2001). The Acquisition of Additive "Focus Particles" in German. In *BUCLD 25: Proceeding of the 25th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, 554-565.

Paterson, K.; Liversedge, S.; Rowland, C. & Filik, R. (2003). Children's comprehension of sentences with focus particles. *Cognition* 89, 263-294.

Penner, Z.; Tracy, R. & Weissenborn, J. (2000). Where scrambling begins: Triggering object scrambling at the early stage in German and Bernese Swiss German. In C. Hamann & S. Powers (eds.) *The acquisition of scrambling and cliticization*. Dordrecht: Kluwer, 127-164.

Recognition of the Prosodic Focus position in German-learning Infants

*Michaela Schmitz¹ Barbara Höhle¹ Frauke Berger¹
Anja Müller² Jürgen Weissenborn²*

¹ University of Potsdam, ² Humboldt University of Berlin

mschmitz@ling.uni-potsdam.de hoehle@uni-potsdam.de

1 Introduction

One way to highlight the focus of a sentence is to mark it prosodically. It is assumed that in sentences with broad focus the nuclear accent universally falls on the rightmost element of the prosodic structure (e.g. Nespor & Guasti, 2002). This may thus be considered the unmarked case of focus-to-stress alignment (FSA). But languages differ in the way narrow focus can be prosodically marked: either the nuclear accent is variable and can be assigned to the focused element in its syntactic position (“stress movement”), or, maintaining the unmarked FSA, the focused constituent is placed at the right edge of the prosodic structure (“syntactic movement”). The languages of the former type (Type A) include German and English, languages of the latter type (Type B) include Italian and Spanish. These two types respectively correspond to relatively fixed and free word-order languages (Donati & Nespor, 2003).

The aim of our study was to find out when and how German-learning children find out to which type their language belongs? Therefore we wanted to investigate whether German-learning children would treat sentences with initial and final stressed narrow focus differently. If there was a bias for one or the other position that would mean that later in development the child would have to overcome this bias in order to realise that there are multiple focus positions in German. Only then prosodic prominence could be used as a reliable indicator of

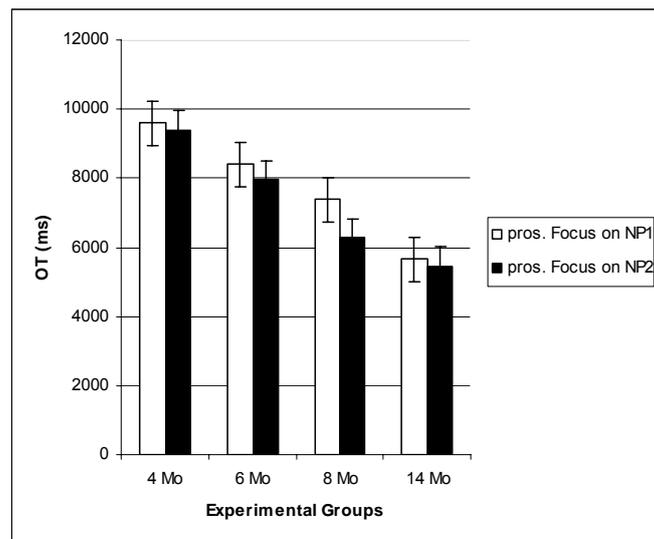
the (semantic) focus of the sentence. Two possible reasons for why the child may show an initial bias for final nuclear stress are, first, that it is the most frequent narrow focus position in German, and second, that it coincides with the universally preferred position for nuclear stress in the case of broad focus.

On the basis of recent findings that children between 6-12 weeks are able to discriminate between languages which differ with respect to whether prosodic prominence falls systematically on the right or the left edge of the phonological phrase (Christophe *et al.*, 2003), we decided to start our experiment at the age of 4 months, repeating it with 6, 8, and 14 month-olds.

2 The Study

In our series of experiments we used the head-turn preference paradigm. The auditory stimuli we presented consisted of 56 canonical sentences in which either the subject NP or the object NP was prosodically highlighted: “Der ENKEL hat Flöte gespielt.” vs. “Der Enkel hat FLÖTE gespielt.”⁶ The sentences were presented in blocks of 7 sentences each, all in all we had 16 experimental trials. The dependent variable was the amount of time the infant spent looking towards each type of stimulus (orientation time). The orientation time towards each condition was calculated and the results of all experiments can be seen in the following graph.

⁶ The capital letters are used to indicate prosodic focus.
<http://www.sfb632.uni-potsdam.de>



The results of the four age groups show an inverted U-shaped curve. The 4-month-olds did not prefer either of the conditions. The 6-month-olds show a tendency to prefer the stimuli in which the first NP is prosodically highlighted. This pattern is manifest in the 8-month-olds, which now have a significant preference ($p = 0.005$) for first NP focus. The 14-month-olds again do not react differently to the two prosodic conditions. When splitting the group of 14-month-olds according to their use of one-word utterances, we find a somewhat differentiated pattern. The group of infants who are still in the pre-word stage show the same response pattern as the 8-month-olds, whereas infants already at the one-word stage make no distinction between the two focus conditions.

The results of the experimental groups can be interpreted as indicating that the 4-month-olds either do not perceive the different prosodic patterns, which is unlikely, as infants already at birth use rhythmic patterns to distinguish their native language from a different language (Mehler *et al.*, 1988) and given the findings of Christophe *et al.* (2003) mentioned above. In the light of the results for the 6-month-olds, who show a tendency and the 8-month-olds who clearly

distinguish between the two prosodic patterns by listening longer to the more infrequent one, a more plausible explanation would be that the 4 month-olds recognize the two patterns but have not yet computed the difference in the frequency of occurrence between the two patterns. The results of the 14-month-old word-users could suggest that during the second year of life German learning infants may already have acquired knowledge that their target language is of Type A in which various focus positions are allowed, possibly on the basis of a beginning lexical and syntactic processing of the elements in prosodic phrase under nuclear stress, (Höhle & Weissenborn, 2000; Shady, Gerken & Jusczyk, 1995). This hypothesis can only be verified by conducting cross-linguistic studies with infants learning a language with a fixed focus position.

Summarizing we would like to propose, that our findings may reflect, like in other areas of prelinguistic language development, the development from a predominantly prosodically driven processing of the input to a processing where prosody interacts more and more with the growing lexical and syntactic knowledge of the child.

References

- Christophe, A., Nespors, M., Guasti, M. T. & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science*, 6(2), 211-220.
- Donati, C. & Nespors, M. (2003). From Focus to Syntax. *Lingua*, 113, 1119-1142.
- Höhle, B. & Weissenborn, J. (2000). The Origins of Syntactic Knowledge: Recognition of determiners in one year old German children. In S. C. Howell, S. A. Fish & T. Keith-Lucas (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 418-429). Somerville, MA.: Cascadilla Press.

-
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143-178.
- Nespor, M. & Guasti, M. T. (2002). Focus-Stress Alignment and its Consequences for Acquisition. *Lingue e Linguaggio*, 1, 79-106.
- Shady, M., Gerken, L. & Jusczyk, P. W. (1995). Some evidence of sensitivity to prosody and word order in ten-month-olds. In D. MacLaughlin & S. McEwan (Eds.), *Proceedings of the 19th Annual Boston Conference on Language Development* (pp. 553-562). Somerville, MA: Cascadilla.

The Influence of Focus Markers on Word Processing and Word Recall in a Second Language

Anke Sennema Ruben van de Vijver Anne Zimmer-Stahl
University of Potsdam

sennema@rz.uni-potsdam.de vijver@rz.uni-potsdam.de

1 Introduction

When adults listen to a foreign language, what makes them turn their attention to certain words, and what makes it easier for them to remember these words? The present study investigated the influence of focus marked by pitch accent, syntax and lexical markers on word processing and word recall in German learners of English, with the aim of examining the use of perceptual strategies for processing focus structure in the second language (L2).

Accent and focus have been found to facilitate word processing in first language processing (L1) (Cutler, 1976). Comprehension also seemed to be facilitated by rapid identification of focussed information as listeners show a processing advantage for words focussed by a preceding question (Cutler & Fodor, 1979). Recently, effects of predicted accent and of focus marked by accent were established for non-native listening (Akker & Cutler, 2003), showing that L2 learners were able to exploit both prosodic structure and sentence semantics. The present study investigated how prosodic, syntactic and lexical marking affected word processing and word recall in non-native listening.

2 Experiments

Experiment I examined the processing of words which are prosodically marked for focus by pitch accent. 60 native German learners of English listened to 40 sentences in three focus conditions (broad focus, narrow focus on the target, narrow focus on another constituent of the sentence) in German (L1) and in English (L2). The length of the target word (1-syllabled or 2-3-syll. words) and its position in the sentence (initial, medial or final) were varied to evaluate a possible effect on word vs. sentence level (for phonetic analyses of the material of Experiment I and II, see van de Vijver et al., 2006). In a closed-set word recognition task, latencies and accuracy of word recognition were recorded. The means of the results are shown in Table 1.

Table 1. Mean reaction times (ms) and mean correct recognition (%) in Experiment I, per language and focus condition

	Language task	Broad focus	Narrow focus	
			on target	not on target
Mean reaction time (ms)	German	1701,4	1694,8	1624,1
	English	1792,4	1755,5	1751,3
Correct recognition (mean %)	German	93,1%	93,7%	92,9%
	English	82,6%	88,9%	85,7%

Focus condition had no significant effect on the processing speed in both the listeners' L1 and L2. ANOVAs carried out on the accuracy scores showed that narrow focus on the target facilitated word recall in the non-native listening task. Narrow focus on the target yielded a more accurate recognition of targets occurring in final position in German, but not in English. In the condition with Narrow focus not on the target, recognition of items in final position was

significantly better in both German and English. Target length itself was in neither language a significant factor, but interacted with position only in two of the German conditions (with narrow focus on/not on target). Thus, although prosody doesn't seem to influence the speed of L2 processing (focus effect), it can still facilitate word recognition in the L2 (position effect).

Experiment II investigated the effect of syntactic cues of focus marking in L2 processing. Focus is often signaled by use of a marked structure, e.g. cleft sentences place a focused noun phrase in post-copular position, indicating that it is new information and thus calling attention to it. Does focus marking by cleft lead to faster word recognition and better word recall in L2 processing? 80 native German learners of English listened to 40 sentences (+/- cleft) in German and English, distributed over two context conditions (+/- question) in order to evaluate the processing advantage for words focussed by a preceding question. A phoneme detection paradigm was used to measure word processing speed; a recall test (4AFC) assessed the representation of the targets in the memory. Means of the results are shown in Table 2. ANOVAs indicated no difference between the language tasks with regard to reaction time but there was an effect for word recall in that English items were more accurately recalled than German ones. Context lead to faster processing only in German and context had in neither language an effect on word recall. Across context conditions, focusing by cleft structure yielded no faster processing in both language conditions. Word recall in the English task was better for items occurring in non-clefted sentences than for targets in clefted sentences and there was no effect of cleft structure on word recall in German. The processing of English words was only faster in clefted than in non-clefted structures if the cleft sentence was induced by a question.

Table 2. Mean reaction times (ms) and mean correct word recall (%) in Experiment 2, per language and syntactic structure, for condition +/- context

	Language	No context		With context	
		cleft	non-cleft	cleft	non-cleft
Mean reaction time (ms)	German	1157,9	1267,6	939,4	1067,8
	English	1318,0	1335,8	1054,9	1220,4
Correct word recall (mean %)	German	49,0 %	44,6 %	42,3 %	42,8 %
	English	50,5 %	57,2 %	48,8 %	53,8 %

In our next experiment (Experiment III, in progress) we will investigate the influence of lexical focus markers on L2 word processing and word recall.

3 Conclusions

Results of Experiment I indicate that prosodic focus in non-native listening facilitates word recall only for narrow focus and that it does not influence the speed of language processing. The overall lack of focus effect might be due to an overshadowing effect of word position. In Experiment II, an advantage of focusing by a question was confirmed only for processing time in native listening: more speech input did not seem to generally help the L2 learners to accomplish the tasks. Syntactic marking did not seem to improve performance in either the L1 or L2. In fact, marked structures such as cleft sentences seem to impede rather than facilitate word recall in the L2. It could be that the complexity of such structures draws too much on the learners' processing resources.

References

- Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition*, 6(2), 81-96.
- Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, 20, 55-60.
- Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition*, 7(1), 49-59.
- van de Vijver, R., Sennema, A., & Zimmer-Stahl, A. (2006). *An analysis of pitch and duration of experimental material*. Paper presented at the International Conference on Information Structure, University of Potsdam.

An Analysis of Pitch and Duration of Experimental Material

Ruben van de Vijver Anke Sennema Anne Zimmer–Stahl

University of Potsdam

vijver@rz.uni-potsdam.de sennema@rz.uni-potsdam.de

1 Introduction

Focus marked by accent has been shown to speed up processing in a native language. It has not yet been investigated whether such an effect is also found in processing a non-native language and whether other means of marking focus have the same effect. To this end we have conducted a number of experiments (see Sennema et al.). The results of the measurement of duration and the pitch of the target words of the experimental material are reported in this paper.

Both pitch and duration are boosted by accent, but being in a clefted constituent or in the scope of a lexical marker do not give an extra boost.

This paper is not intended as an analysis of focus and its markers in German and English: There is only one speaker per experimental set, which makes a generalization to the German or English population impossible. Moreover, the material has been controlled for prosodically, but not segmentally.

It is nevertheless important to know the phonetic properties of the material of any auditory linguistic experiment, since (sorry for stating the obvious) they are a factor in the experiment.

2 Material

There are three sets of experimental material. In the first set the target word is prosodically marked for focus, by means of an accent. There is only an English version of this material.

The second set consists of material in which the target word is syntactically marked for focus. The target word is in a clefted constituent. There is an English and a German version.

The third set, finally, consists of material in which the target word is in the scope of a lexical marker for focus. There is an English and a German version.

In all sets the duration, the lowest and the highest pitch of the first and the second syllable of the target word have been measured. The pitch values have been converted from Hz to ERB.⁷

3 Prosodic means of marking focus

This material was read by a female native speaker of American English. The material consisted of question–answer pairs. In the answer either the target, a bird name, was accented or an adjective preceding the bird name. This accent was induced by the preceding question.

- (1) Prosodic marking of focus .
- a. What noisy animal did some rude children blame the ruckus on?
Some rude children blamed the noisy GAPPET for the ruckus.
 - b. What kind of gappet did some rude children blame the ruckus on?
Some rude children blamed the NOISY gappet for the ruckus.

There were 40 such pairs. Only the analyses of the answer sentences are presented here, since these were the sentences the subjects were tested on. This

⁷ Equivalent Rectangular Bandwidth:

$$\text{ERB} = 16.6 \times \log\left(1 + \frac{H_x}{165.4}\right) \text{ (Traunmüller, 1990).}$$

study was used as a pilot to test whether the phoneme monitoring paradigm was suited for our purposes and therefore we only recorded an English version. For all other experiments there are always two versions, an English version and a German version.

4 Syntactic means of marking focus

In the next set of experiments the target word appeared either in a cleft structure or in a default declarative sentence. The target was again accented itself, or preceded by an accented adjective. The English sentences were read by a male native speaker of British English; the German sentences were read by a native speaker of German.

(2) Syntactic marking of focus (English)

- a. It's the stale GANNET that is suffering from city development.
- b. It's the STALE gannet that is suffering from city development.

(3) Syntactic marking of focus (German)

- a. Es ist der faule KABU, der stundenlang auf einen Fuß steht.
- b. Es ist der FAULE kabu, der stundenlang auf einen Fuß steht.

5 Lexical means of marking focus

The target word in this set is within the scope of a lexical marker of focus (*even* or *only* in the English version). This experiment is in its preparatorian stage and only the English material was available for analysis.

Lexical means of marking focus What kind of animal did an ill lawyer move onto the sidewalk?

- a. An ill lawyer moved a RUTHLESS ganta onto the sidewalk.
- b. An ill lawyer moved only a RUTHLESS ganta onto the sidewalk.
What ruthless animal did an ill lawyer move onto the sidewalk?
- c. An ill lawyer moved a ruthless GANTA onto the sidewalk.
- d. An ill lawyer moved only a ruthless GANTA onto the sidewalk.

6 Discussion

In this paper, we reported on measurements of duration and pitch of target words of three sets of our experimental material. We wanted to investigate whether non-prosodic ways of marking focus (syntactic and lexical) had an effect on the prosody of our target words. It turned out that it did not. Accent boosts both duration and pitch, but neither cleft, as a way of syntactically marking focus, nor lexical focus markers had an additional effect. This is certainly important for the interpretation of our results.

References

- Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, 88, 97–100.

Processing Information Structure

Shravan Vasishth

University of Potsdam

vasishth@acm.org

1 Main goals

Information structure is concerned with all devices that restructure linguistic information in order to optimize information transfer in discourse. By definition, information structure is intimately related to the real-time access and integration of information—in other words, to language processing in real time. Although there are several detailed models of sentence comprehension, some of them even implemented, remarkably there exists not a single, implemented model of language processing that takes information structure constraints (such as givenness, word order, and prosodic structure) into account. The principal goal of this project is to develop such an implemented process model.

In the initial stage of this project, we focus on the notion of ‘prominence’ or ‘salience’ in information structure and its relationship to activation of items in memory. Although one end-product of information-structuring (such as focus and topic) is increased prominence of linguistic elements, few attempts (cf. Almor 2000, Arnold 1998, Kintsch 1988) have been made to connect these ideas to independently operating cognitive constraints on activation (such as decay, associative retrieval interference) that derive from a vast literature in psycholinguistics and cognitive psychology. Since the object of inquiry (identifying and explaining the constraints on linguistic objects) is identical in psycholinguistic and information-structure research, establishing such a connection is vital.

The long-term goal of this project is to deliver a comprehensive, implemented computational model of human sentence comprehension that integrates constraints and processing cues imposed on language by information structure. This goal will be achieved through computational modeling as well experimental research involving several languages.

2 Understanding prominence: experimental investigations

There is a strong claim in the psycholinguistic literature that prominence (as determined by, for example, pronoun resolution) is affected by all available cues including information structure (Kaiser & Trueswell, in press). There are, however, claims in the literature that pronoun resolution is unaffected by information structure in certain languages. One such claim has been forwarded for Hindi (Prasad 2003). If true, this claim has important consequences for a model of processing that incorporates information structure cues: there would be no language-independent parsing mechanism, rather, comprehension processes in Hindi would remain relatively unaffected by IS while a language like German would be sensitive to IS cues. Since a prerequisite for the modeling task is deciding whether a language-independent parsing strategy be adopted, we begin with a series of four offline and eyetracking experiments, described in detail in the poster and proposal, to first identify if this claim is true. Three other experiments are planned with German, which is known to be sensitive to information structure for pronoun resolution. In addition, we also explore experimentally whether antecedent resolution as a signal of prominence can be affected adversely by IS-independent working memory constraints such as similarity-based interference.

3 Computational modeling

Several experimental results have shown that pronoun resolution processes are affected by (a) first-mention (Gernsbacher 1989), (b) recency (Grosz 1977), and (c) givenness, syntactic position, prosodic emphasis and/or focus-status of antecedents (Ariel 1990, Arnold 1998). Pronoun resolution can also slow down when multiple antecedent-candidates are present (Fredriksen 1981). Interestingly, all these constraints, which have hitherto been stated as facts about prominence without any underlying explanation, fall out of independently motivated principles of working memory. These key ingredients from working memory are (a) decay (b) reactivation through repeated use (c) associative retrieval interference (d) cue-based retrieval (Anderson et al. 2004). An existing sentence processing model (Lewis & Vasishth 2005, Vasishth & Lewis 2004, 2006) utilizes precisely these constraints and is therefore well-suited for modeling prominence as antecedent resolution. In the project, we will extend this computational model to explain existing experimental results on antecedent resolution difficulty.

In addition to modeling existing experimental results, the experimental work outlined in Section 2 will feed into the modeling work. Specifically, if the experimental results show that information structure affects pronoun resolution in a language like Hindi (for which it has been claimed that IS plays no role), the model must then incorporate a language-independent process of incremental information structure construction. Furthermore, if similarity-based interference and information structure cue can be shown to interact in determining prominence, this will be a strong validation of the central assumption in this project that activation, as defined in psycholinguistic research, is the underlying explanation for prominence: both similarity-based interference and information structure will then have the effect of damping or increasing activation. All

constraints, be they cognitive/working memory related or information structural, will ultimately impact a single explanatory variable, activation of items in memory. Since a detailed process model of sentence parsing based on constraints on activation already exists (Lewis & Vasishth 2005), adding a theory of information structure processing that also affects the constraints on activation will result in a model of language comprehension that receives independent support from cognitive psychology as well as linguistic theory.

References

- Almor, Amit (2000). Constraints and mechanisms in theories of anaphor resolution. In: M. Pickering et al., *Architectures and mechanisms for language processing*.
- Arnold, J.E. (1998). *Accessing noun-phrase antecedents*. London: Routledge.
- Kintsch, W. (1988). *Comprehension: A paradigm for cognition*. CUP: New York.
- Lewis, R.L & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 1-45.

From Data to Insights: Exploiting Linguistic Data

Stefanie Dipper Michael Götze Julia Ritz Stavros Skopeteas

University of Potsdam

{dipper,goetze,julia}@ling.uni-potsdam.de

In our talk, we focus on the role of the linguistic data that is being collected and exploited within the SFB. One of the main interests of the SFB is to work out in detail the different factors that contribute to Information Structure (IS) in different languages, and clarify the way these factors interact. To this end, several projects have collected and researched data from typologically diverse languages, such as Prinmi, Georgian, or Yucatec Maya.

Collecting data The linguistic data that serves as the basis of this research has to satisfy (at least) two constraints: it has to provide evidence for all putative IS-related factors and interactions between them, and it must allow for comparative evaluations of the data across the languages. To gain such data, project D2⁸ has developed a *questionnaire* with different types of experiments, like map tasks or scenario descriptions that are presented visually. The experiments are designed to yield a complete picture of IS factors. For instance, given the assumption that givenness and theta-roles are such relevant factors, a series of experiments tests for all possible combinations of these factors: in the first experiment, the agent solicited by the question is new in the context whereas the patient is given information, i.e., part of the question's background; in the second experiment, the patient is new and the agent is given; etc.

⁸ Project D2 (Caroline Féry, Gisbert Fanselow, Manfred Krifka, Stavros Skopeteas, Sam Hellmuth, Ines Fiedler, Anne Schwarz): *Typology of Information Structure*

The questionnaire has been used by trained native speakers, from about 15 different languages, to elicit data from their respective languages. As a result, it is now possible to directly compare the means that these different languages deploy to verbalize the same situations, like the ones describe above.

Annotating data To maximize profit from the data elicited by the questionnaire, the data is being annotated with various types of linguistic information, including, e.g., morpheme segmentation and gloss, prosodic properties, basic constituent structures (including the annotation of syntactic categories, grammatical functions and theta-roles), information status, topic, and focus. The annotations thus provide access to complex, interesting phenomena, such as the interplay of givenness and theta-roles.

To make the annotations consistent and comparable across the languages, SFB-wide groups have developed *annotation guidelines* for all layers that are to be annotated. The guidelines have been used by the native speakers to annotate the data they elicited. Similarly, the guidelines serve as an interpretation guide to all those who want to exploit the data and its annotations in their research.



Exploiting data The data and its annotations are stored and maintained in the database *ANNIS*⁹, developed by project D1¹⁰. Prior to import into the database, the data is converted to an interchange format, *PAULA*¹¹. *PAULA* is an XML-based, *stand-off* representation format that allows for homogeneous

⁹ *ANNIS: A linguistic dabase for Annotated Information Structure*

¹⁰ Project D1 (Manfred Stede, Stefanie Dipper, Michael Götze, Julia Ritz): *Linguistic Database for Information Structure: Annotation and Retrieval*

¹¹ *PAULA: Potsdamer Austauschformat für linguistische Annotation* (Potsdam Interchange Format for Linguistic Annotation)

<http://www.sfb632.uni-potsdam.de>

representation of data that has been annotated by means of different annotation tools. It provides data structures for tier-based annotations, such as phonological phrases, as well as graph- or tree-based structures, such as syntactic or discourse trees.

The database *ANNIS* allows the user to browse the data via a web-based graphical user interface and display the annotations, in order to develop linguistic hypotheses. Next, the user can systematically search for data confirming or refuting the hypotheses, by means of the query facilities provided by *ANNIS*. For instance, the user can find out which of the languages, in which contexts, violate the general principle that given information precedes new information. Another research question could ask for the use of presentational cleft sentences in “all new” contexts; according to the data in *ANNIS*, this is a rather frequently-used option in English (43 clefts vs. 60 non-clefts) but no option in Prinmi (0 clefts vs. 76 non-clefts).

Finally, the data can be used for *feature mining*, in that feature co-occurrences in the annotations are automatically computed and correlations are presented to the user. Such correlations may hint at interesting phenomena, such as unexpected dependencies between certain features. Similarly, the technique can be used to systematically search for putative annotations errors, thus promoting quality assurance.

Sustainability of Linguistic Data

Christian Chiarcos

Potsdam University

chiarcos@uni-potsdam.de

1 Avoiding data graveyards

The poster presents a new research initiative addressing the issue of sustainability of linguistic resources. The initiative is a cooperation between three collaborative research centres in Germany – SFB 441 “Linguistic Data Structures” in Tübingen, SFB 538 “Multilingualism” in Hamburg, and SFB 632 “Information Structure” in Potsdam/Berlin (Schmidt et al. 2006).

Researchers often create linguistic data with a specific linguistic theory and a concrete research question in mind. Data formats as well as the tools used to create, edit and analyse corpora are tailored to the specific task at hand, and little attention is paid to the question of how these corpora could be exchanged or reused for other purposes in the future. Often, the resulting data is dependent on a single piece of software, so it becomes difficult to use when this software is no longer supported by its developers. Even where no such fundamental technical obstacles exist, the lack of proper documentation or difficulties in adapting a resource to the requirements of a new research question can greatly hamper data exchange and reuse.

The aim of the project is to develop methods for sustainable archiving of the diverse bodies of linguistic data used at the three sites. The primary goal of our initiative is to convert the data collected by the three collaborative research centres into a comprehensive and sustainable linguistic corpus archive that we aim to be accessible and usable by researchers and applications for at least five

decades. In addition, methodologies and rules of best practice for data storage, annotation, and access will be developed. We see our work as a kind of blueprint for comparable initiatives.

2 Aspects of sustainability of linguistic data

In the last two decades, the amount of language data collected for linguistic research purposes has increased dramatically. Usually, data formats and annotation standards as well as the content depend on the research questions a specific project pursues. In conjunction with technological changes, this diversity causes a high degree of heterogeneity, thus, it can be difficult to exchange these data collections with other groups or to reuse them in different research contexts when the initial project is completed.

The three centres conjoined in our initiative have collected language data over a period of several years and have processed it according to their specific research questions. Thus, the data collected is extremely heterogeneous. We have identified three fields with seven main areas of future work:

technical aspects

(i) development of data formats, (ii) means for data distribution and data access, and (iii) query interfaces

maintenance

(iv) data integration, and (v) integration of linguistic terminology

rules of best practice

(vi) documentation, and (vii) legal aspects of linguistic data collections

Here, technical and maintenance aspects are discussed with greater level of detail.

3 Technical Aspects of Sustainability

Just as the SFB-specific data formats (TUSNELDA, EXMARaLDA and PAULA) are set out to generalise over a number of project-specific data models and formats, one of our primary goals is to develop a **generalising data model** which generalises over these. An XML-based format such as the one we are aiming at seems to be a promising candidate for achieving long-term usability for digital data. We are currently exploring to what extent the NITE Object Model (Carletta et al. 2003) can be used as a starting point for a sustainable data format.

Once linguistic resources are available in a form that makes them suitable for reuse in other research contexts, **archiving and disseminating methods** have to be developed. Short and medium term data dissemination has to focus on methods that allow researchers quickly to discover an existing resource and to assess its relevance for their research purposes. For these aspects, web interfaces are the most promising approach as they offer access to corpus documentation and allow for querying corpus metadata. We therefore plan to develop both web-based data distribution methods and digital versions of corpora to be distributed on offline media such as DVDs.

Data distribution methods will have to provide **means of querying** metadata and linguistic data. In terms of sustainability, it is crucial that such query mechanisms are intuitively usable by a wide range of researchers, as the usability of interfaces has a decisive impact on whether and how a resource can be reused. Moreover, it is desirable that similar query tasks are approachable in a similar manner across heterogeneous data sets. The solutions developed at the

three sites already provide a number of query mechanisms for the respective data models. We will build upon this work, adapting and extending these solutions to the data format to be developed.

4 Sustainable Maintainance

As long as a data resource is only used within the project that created it, often only minimal attention is paid to systematically and explicitly recording **metadata** (information about the composition of the corpus, details about speakers or authors, details on annotation levels, etc.), because the researchers have been involved directly in the creation of the corpus and therefore have such information “in their heads”. However, when the resource is to be made available to other researchers, it becomes crucial for such information to be represented in a systematised, digital form. Thus, our joint initiative plans to compile a comprehensive set of metadata.

Furthermore, different **terminologies** applied during the annotation of the corpora have to be integrated into one overarching and well-defined terminological back-bone which can be used for theory-neutral tag definition and tagset-neutral corpus querying. Annotation details such as definitions of tags, tag names, labels of syntactic structures, etc., can be highly idiosyncratic or restricted to the conventions of a specific community. Moreover, tagset descriptions and annotation guidelines occasionally lack definitions at all. Currently, we are developing an ontology inspired by the EAGLES recommendations (Leech & Wilson 1996). Similar to the E-MELD project (Farrar & Langendoen 2003), individual tagsets will be linked to ontological concepts to allow for tagset-neutral querying within a well-defined theoretical framework.

References

- Carletta, Jean, Jonathan Kilgour, Timothy O'Donnell, Stefan Evert, Holger Voormann (2003): The NITE Object Model Library for Handling Structured Linguistic Annotation on Multimodal Data Sets. In: *Proceedings of the EACL Workshop on Language Technology and the Semantic*.
- Farrar, Scott and D. Terence Langendoen (2003). A Linguistic Ontology for the Semantic Web. *GLOT International* 7 (3), pp. 97–100.
- Leech, G. and Wilson, A. (1996), EAGLES Recommendations for the Morphosyntactic Annotation of Corpora,
<http://www.ilc.cnr.it/EAGLES/annotate/annotate.html>.
- Schmidt, Th., Ch. Chiarcos, T. Lehmberg, G. Rehm, A. Witt, E. Hinrichs (2006), Avoiding Data Graveyards: From Heterogeneous Data Collected in Multiple Research Projects to Sustainable Linguistic Resources, paper to be presented at the *E-MELD 2006 Workshop on Digital Language Documentation: Tools and Standards. The State of the Art*. Ypsilanti, June 2006.

ANNIS: A Linguistic Database for Complex Multilevel Annotation

Manfred Stede Stefanie Dipper Michael Götze Julia Ritz

University of Potsdam

{stede,dipper,goetze,julia}@ling.uni-potsdam.de

We present *ANNIS*¹², a linguistic database that aims at facilitating the process of exploiting richly annotated language data. In particular, *ANNIS* provides suitable means for visualizing and querying complex multilevel annotations.

ANNIS is being developed in the Collaborative Research Center SFB 632 on Information Structure, in which data of various languages is collected and annotated at multiple levels – levels that contribute in ways yet to be determined to the information structural partitioning of discourse and utterances.

Annotated data in the research center varies with respect to size (single sentences vs. narrations), modality (monologue vs. dialogue, text vs. speech) and language. Furthermore, annotations use different data structures (attribute-value pairs, trees, pointers, etc.) and stem from different specialized tools (such as *EXMARaLDA*, *annotate*, *RSTTool*, and *MMAx*). Currently, creators and users of this data are linguists from diverse linguistic fields and traditions such as historical linguistics, african studies, syntax or typology, with little or no programming skills. Once the data is published, the heterogeneity of its users will increase further. Thus, besides *Data heterogeneity* also *Accessibility* for different users is an important requirement for a corpus exploration tool in our research scenario. Finally, the tasks connected to this

¹² <http://www.sfb632.uni-potsdam.de/annis>

language data vary - besides manual inspection, these include statistical analysis and further or re-annotation. Thus, aspects of *Data Reuse* have to be considered.

With *ANNIS*, we approach these requirements as follows. Data annotated with the specialized tools is converted into *PAULA* („Potsdamer Austauschformat für Linguistische Annotationen“), a generic standoff data interchange format (Dipper 2005), which serves as an input format for *ANNIS*.

ANNIS itself is a Java servlet application that can be accessed with standard web browsers. In its current state, it is not database-backed; data is read into memory and exploited for querying and visualization in memory.

The system supports the concurrent visualization of different types of annotation, where the user can directly customize the appearance by selecting the annotation to be displayed. In addition, a user-adaptable format file specifies features such as coloring or ordering of the annotation layers.

The *ANNIS* query language offers a rich set of query operators, enabling the evaluation of queries across multiple annotation layers with temporal (e.g. precedence or overlap) and hierarchical (e.g. dominance or sibling) relations. Regular expressions and wildcards are supported.

In our presentation, we illustrate the use of *ANNIS* with research questions from three different linguistic areas. With data from the Old High German translation of Tatian (project B4: *The role of information structure in the development of word order regularities in Germanic*), we exemplify the use of temporal queries in multilayer annotations for research in historical linguistics. Data elicited with the Questionnaire on Information Structure (*D2: Typology of Information Structure*, cf. Götze et al. to appear) and annotated with information from various linguistic levels are of particular interest for typological studies. We show how *ANNIS* facilitates the search across data of typologically diverse languages. The *Potsdam Commentary Corpus* (Stede 2004) with annotations of

morphosyntax, coreference, discourse structure and information structure is a prototypical example for complex multilevel annotated corpora. We will illustrate the capabilities of the ANNIS query languages for the search in hierarchical data structures, investigating relations of sentence and discourse level phenomena.

References

- Dipper, Stefanie. 2005. XML-based Stand-off Representation and Exploitation of Multi-Level Linguistic Annotation. In *Proceedings of Berliner XML Tage 2005 (BXML 2005)*, 39-50, Berlin, Germany.
- Dipper, Stefanie, Michael Götze, Manfred Stede, and Tillmann Wegst (2004). ANNIS: A Linguistic Database for Exploring Information Structure. In Ishihara, S., M. Schmitz and A. Schwarz (eds.): *Interdisciplinary Studies on Information Structure 1 (ISIS)*, 245-279, Potsdam, Germany.
- Götze, Michael, Stavros Skopeteas, Thorsten Roloff and Ruben Stoel (to appear). Towards an Exploration Infrastructure for a Cross-Linguistic Production Data Corpus. In *Proceedings of the Sixth International Tbilisi Symposium on Language, Logic and Computation*, Batumi, Georgia.
- Stede, Manfred. 2004. The Potsdam Commentar Corpus. In *Proceedings of the ACL Workshop on Discourse Annotation*, 96-102, Barcelona, Catalonia.

Towards a Typology of Information Structure: Insights From Cross-linguistic Data Collection

Sam Hellmuth Stavros Skopeteas

University of Potsdam

{shellmut, skopetea}@uni-potsdam.de

1 Introduction

This talk presents insights gained through a cross-linguistic production study being carried out within the SFB by project D2 (G. Fanselow, C. Féry, M. Krifka, S. Hellmuth, S. Skopeteas, A. Schwarz, and I. Fiedler). The target of this project is to develop tools that can be used in linguistic fieldwork to elicit the means of encoding information structure in an object language. For this purpose the project has developed a *Questionnaire on Information Structure* that contains a number of elicitation tools for the collection of a large volume of data for observational research on information structure. The bulk of the questionnaire comprises production experiments that manipulate a number of discourse variables in order to induce sentences with different information structures.

The main contribution of this kind of linguistic evidence is that it allows for insights into the ways that grammatical structure influences or constrains language production in a near-naturalistic experimental context. It is obvious that data gained in this way does not exhibit all structures that could be judged as acceptable on the basis of speakers' intuitions. This data allows however for insights as to the choices that speakers make under certain discourse conditions. In our presentation, we want to illustrate that languages not only differ with respect to the means that are available in the grammar, but also with respect to the structures that speakers choose in particular contexts in order to encode

information structure. Even when a particular strategy of encoding information structure is available in each of their different grammars, languages may differ with respect to the choice to use this strategy or not, and this difference depends on the availability of alternative means to fulfill the same purpose.

We illustrate this point on the basis of three different discourse conditions, as implemented in different experiments of the QUIS:

- one argument is new information, the other argument is given;
- one argument is new information, solicited by a question, the rest is given;
- one argument is exhaustively identified, the rest is given.

Data are presented primarily from English, Georgian, Greek, and Yucatec Maya.

2 Information structure strategies

2.1 Cleft constructions

We know that the grammars of English, Georgian, and Greek allow the formation of identificational cleft sentences. In the discourse condition in which one argument is exhaustively identified however, we observe that only English speakers produce cleft sentences. Speakers of the other languages use other means to encode the same function (prosody in Greek, prosody and change of word order in Georgian, movement to the focus position in Mayan).

2.2 Word order

The four object languages have different word orders: English is an SVO language that does not allow for movement, Greek is an SVO/VSO language that employs movement to encode topicalization, Georgian is a scrambling SOV

language and Yucatec Maya is a VOS language that allows for movement to information structurally designated positions in the preverbal domain. Our production study shows that languages that allow for non-canonical word orders do not make use of them in the same way. When the agent is new information and the patient is given, Georgian speakers choose OSV/OVS orders, while Greek speakers do not use this option at all, even though it would be ‘grammatical’.

2.3 Passivization

The grammars of the object languages differ with respect to the availability of passivization. English and Yucatec Maya are known to have productive passivization. Greek verbs form the passive voice through inflection, displaying all the idiosyncracies of inflectional morphology: in particular, medio-passive forms of verbs of action and spatial transposition are mainly interpreted as middles (though a passive reading is possible). Georgian has productive passive morphology that is reported however to only rarely occur in spontaneous discourse. Our experimental study shows that languages differ with respect to the choice of a passive construction, depending on the available alternatives. When the patient is given information and the agent is new, Yucatec Mayan speakers use passive verbs almost exclusively, whereas English speakers choose one of two alternatives: a passive verb with a patient subject or an active verb with the agent as subject and a prosodic strategy to mark the new agent.

2.4 Prosodic strategies

Cross-linguistically there are known to be a range of prosodic reflexes of categories such as new information, givenness and exhaustive focus. In English for example, the distinction between new and given items can be marked by ‘de-accenting’ of the given item (Ladd 1996) and a corrective focus can

(arguably) be expressed by means of a special pitch accent on the focussed word or by insertion of a phrase boundary after it (Selkirk 2002). In Greek a special pitch accent marks narrow focus, and is accompanied by post-focal de-accenting (Arvaniti & Baltazani 2005). The prosodic reflexes of information structure in Georgian is under-researched and an initial survey reported here seeks to establish what prosodic strategies are used if any. In contrast, Yucatec Maya is a tone language, thus all information structure strategies are expected to be non-prosodic; this expectation is evaluated against the facts of the elicited data. In particular it is of interest to establish whether prosodic and syntactic strategies co-occur or are used complementarily by speakers.

3 Conclusion

Collection of real production data from native speakers allows for insights that are not possible from data elicited in the form of speaker-intuitions. Grammars display an array of structures for encoding information structure, but speakers choose from among them which to use. We suggest that these choices are based on the particular functions of each grammatical structure in their language and on the paradigmatic alternatives available to speakers to encode each information structural function.

References

- Skopeteas, S., I. Fiedler, A. Schwarz, R. Stoel, G. Fanselow, C. Féry, & M. Krifka. 2006 in press. Questionnaire on Information Structure (QUIS). *Interdisciplinary Studies on Information Structure*

Information Structure in the Project Languages of D2 ‘Typology of Information Structure’

*Stavros Skopeteas, Gisbert Fanselow, Caroline Féry, Sam Hellmuth,
Manfred Krifka, Ruben Stoel*

University of Potsdam

{skopetea,shellmut}@rz.uni-potsdam.de

1 Introduction

This abstract presents the outline structure of a *family of posters* each of which provides an introduction to the information structure in a particular language from the list in Appendix 1 below. These are the languages which are currently the subject of study in the D2 Project ‘Typology of Information Structure’. The author of each poster is (unless noted) the Associate Field Researcher responsible for data collection using the Questionnaire on Information Structure (Skopeteas, Fiedler, Schwarz, Stoel, Fanselow, Féry & Krifka 2006).

Together these posters demonstrate the extent of work ongoing in the D2 project, and illustrate both the value and potential of a parallel corpus of spoken data, elicited in a known information structure context, in a typologically diverse range of languages.

2 Outline structure of each poster

1) Object language

General information about the language and language situation is given, and an indication of where the language is spoken. The time and place(s) in which data elicitation took place is described; in cases where the field or language situation is unusual additional information is provided.

2) Information Structure

This section of each poster offers a short outline of the fundamental properties of the language with respect to information structure, whether such marking is prosodic, syntactic or morphological (or some combination of all three). In cases where relevant previous work exists, key points from the literature are noted.

3) Empirical observations

The main body of each poster presents one or more representative issues or observations about the object language that have been possible through the data collection gathered using the Questionnaire on Information Structure. Most often there is one issue related to phonology and one to syntax.

For languages that have been already intensively investigated the relation between the collected data and available knowledge about the language are explored: are existing assumptions about the language confirmed, or are additional phenomena revealed? For less studied languages the first findings regarding information structure are reported, together with indications of potentially fruitful avenues for future research.

4) Future work

A final section summarises the poster and outlines plans for further investigation.

Additional languages for which data are being collected, or will be collected shortly, for D2 using the QUIS questionnaire, and for which a first analysis of the data is therefore not yet available, include Mandarin Chinese and Egyptian Arabic.

A printed copy of each poster is available to collect beside the poster stand. Although the Associate Researchers are not able to attend the conference in

person please feel free to address questions to D2 personnel who will be nearby. The D2 personnel are in constant conversation with the Associate Researchers about each stage of data collection, annotation and analysis.

References

Skopeteas, S., I. Fiedler, A. Schwarz, R. Stoel, G. Fanselow, C. Féry, & M. Krifka. 2006 in press. Questionnaire on Information Structure (QUIS). *Interdisciplinary Studies on Information Structure*.

Appendix 1 D2 ‘Typology of Information Structure’ Project Languages

language	language family, <i>where spoken</i>	word order	word prosody	authors
Dutch	Germanic Netherlands	SOV	intonation	Hanneke Van Hoof
Québec French	Romance Québec, Canada	SVO	intonation	Alain Thériault
Maung	Australian N Territory, Australia	SOV /SVO	intonation	Ruth Singer
Georgian	Kartvelian <i>Georgia</i>	SOV	intonation	Rusudan Asatiani
Teribe	Chibchan <i>Panama</i>	SOV	lexical pitch accent	Diego Quesada
Prinmi	Sino-Tibetan China	SOV	lexical tone	Teng Si Chi
Hungarian	Uralic Hungary	SVO	intonation	Krisztián Tronka
Japanese	Japanese Japan	SOV	lexical pitch accent	Yukiko Morimoto
Konkani	Indo-Aryan Maharashtra, India	SOV	intonation	Caroline Menezes
Yucatec Maya	Mayan Mexico	VOS	lexical tone	Frank Kügler (D2), Stavros Skopeteas (D2), Elisabeth Verhoeven
Greek	Indo-European Greece	SVO/VSO	intonation	Thanasis Georgakopoulos, Janis Kostopoulos, Georgios Markopoulos, Stavros Skopeteas
English	Germanic USA, UK etc	SVO	intonation	Sam Hellmuth (D2), Elizabeth Medvedovsky (SFB632), Stavros

				Skopeteas (D2)
German	Germanic Germany, Austria etc	SOV	intonation	Sam Hellmuth (D2), Katarina Moczko (D2), Andreas Pankau (D2), Stavros Skopeteas (D2)

Split-noun Sentences in Mandarin and in German: Production and Perception

Bei Wang Caroline Féry

University of Potsdam

{bjwangbei, caroline.fery}@gmail.com

The aim of this study is to compare the prosodic realization of the split nominal construction in Mandarin and in German (Bücher hat Maria interessante gelesen. ‘Mary has read interesting books,’ literally: books has Mary interesting read), both in topic and focus context. Our second goal is to study the extent to which the split sentences are accepted by listeners and whether the prosodic characteristics are sufficient to distinguish the split sentences in a topic context from those in a focus context.

Pan’s syntactic analysis (2003) suggested that Mandarin split nominal constituent is focused. He argued that when the focus phrase that contains the empty head noun is adjacent to the topicalized noun, it can establish a predicative relation with the topicalized noun. Féry (2005) proposes that the marked tonal pattern of split structures in German triggers an additional prosodic phrase. Furthermore, there might be two information structural patterns in German: the split noun is topicalized and the remainder of the sentence is generally focused, or the split noun is focused and the remainder is a deaccented background.

1 Speech production Experiments

Six female German speakers and 7 Mandarin speakers (5 female and 2 male speakers) participated in the speech production experiments. 20 Mandarin and 8 German target sentences were constructed, as illustrated in (1a) and (1b). For each target sentence, a context question sentence elicited the split noun either as a topic or a focus of the target sentence. (2a) and (2b) are examples for both Mandarin and German context sentences.

Target sentences:

(1a) *shuqian BaoXin diu le san zhang.* (bookmarkers BaoXin lose ASP CLAS. ‘Bao Xin lost three bookmarkers.’)

(1b) *Birnen hat sie einige gekauft.* (pears AUX she some bought. ‘She bought some pears’)

Context question sentences:

(2a) **Focus questions:** *Bao Xin diu le san zhang shenme?* (What did BaoXin lose three of?)

Topic questions: *Wo tingshuo baoxin diu le shuqian he benzi, shi zhen de ma?* (I heard Bao Xin lost bookmarkers and notebooks. Is that true?)

(2b) **Focus questions:** *Was hat sie einige gekauft?* (What did she buy some of?)

Topic questions: *Ich habe gehört, dass sie Birnen und Äpfel gekauft hat. Ist es wahr?* (I’ve heard that she bought pears and apples. Is that true?)

The participants were asked to read the target sentences naturally after hearing the question. The F0 and the boundaries of each syllable (Mandarin)/word (German) were manually labeled for every vocal cycle using Xu’s Praat script (1999, 2005).

Figure 1 and 2 show the intonation contours (10 points of each syllable/word and no pauses) of both the two contexts for Mandarin and German respectively.

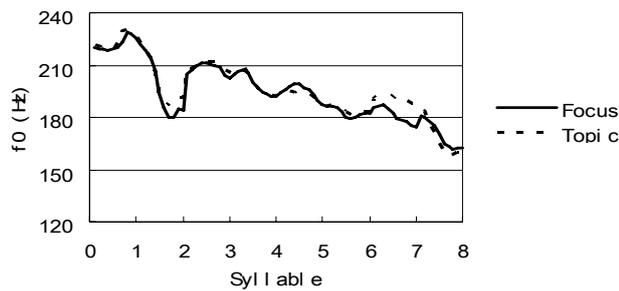


Figure 1. The intonational contours of **Mandarin** split sentences in the two contexts

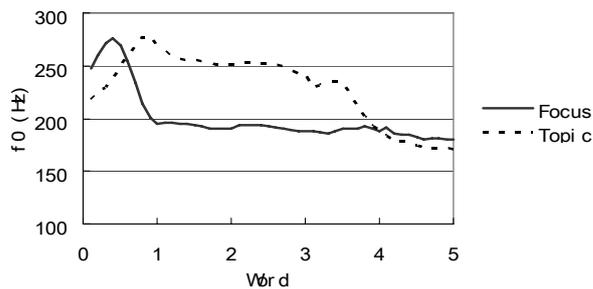


Figure 2. The intonational contours of **German** split sentences in the two contexts

Surprisingly, in Mandarin there is little difference in pitch between the two contexts, especially on the split noun. The modifier, which is at the end of the sentence, is slightly higher in the topic context than that in the focus context.

In German, when the split noun is in focus, it has a falling tone and F0 is compressed afterwards. However, when it is a topic, a hat pattern is the consequence. The split noun has a rising tone and the remaining of the contour declines slowly until the falling tone on the modifier, which is accompanied by a steep fall. In this case, the modifier is the focus of the sentence.

In German, only 3 of the 6 speakers produced pauses after the split noun, while in Mandarin, all the speakers did. The Mandarin speakers' pauses were considerably longer than the German speakers' ones (see table 1). Interestingly, the pause in the topic context was significantly shorter than that in the focus context in both languages.

Table 1. Pause length after the split noun

	German (only 3 speakers)	Mandarin (all 7 speakers)
Focus context	35.8 ms	248 ms
Topic context	5 ms	189 ms

2 Speech Perception Experiments

The question and answer pairs from the speech production experiments were cross-spliced and matched. They were played in sequence to the participants. The task is to judge whether the intonation of the answer was a good match to the question. A 10-point scale was used, with 10 being the best. 21 Mandarin and so far 6 German participants joined in the experiment.

Table 2. The average value of the judgment (10 is the best)

Question \ Answer	Mandarin		German	
	Focus	Topic	Focus	Topic
Focus	6.83	5.91	7.54	4.04
Topic	5.63	5.11	6.5	4.93

The important result in table 2 is that the focus answer is always attributed better scores than the topic answer in both contexts and the distinction is larger in German than in Mandarin. With the topic questions, the matching answers are not better in both languages. This can be an effect of our design, and prompts us to conduct a new set of experiments.

3 Conclusions

Split sentence is a marked structure in Mandarin, that induces a long pause after the split noun. In German, the construction is more natural, and the pause after the noun is much shorter and not obligatory. The difference of the intonational contour between the two contexts in Mandarin shows that in the topic context, the modifier is realized with a raised F₀, as a narrow focus, while in the focus context, it is part of a wide focus. In both contexts, the split noun is accented with high F₀. In German, the split sentence in the topic context shows the typical hat pattern, while in the focus context, the split noun has a falling tone and the remainder of the sentence is deaccented.

As revealed by the perception experiments, listeners perceive the difference quite well of the focus context. The intonational contours differ more between the two contexts in German than in Mandarin, and as a result, the German listeners can distinguish the two patterns better than the Mandarin listeners.

References

- Pan H.H., Hu J.H. (2003) Head Noun Movement, Focus, and Topicalization in Mandarin Chinese, In: Xu, J. Donghong Ji, and Kim Teng Lau (eds.) Chinese Syntax and Semantics, Singapore, 2003,119-156.

Tone, Intrinsic f0, and Intonational Focus Marking in Chichewa

Bernd Pompino-Marschall

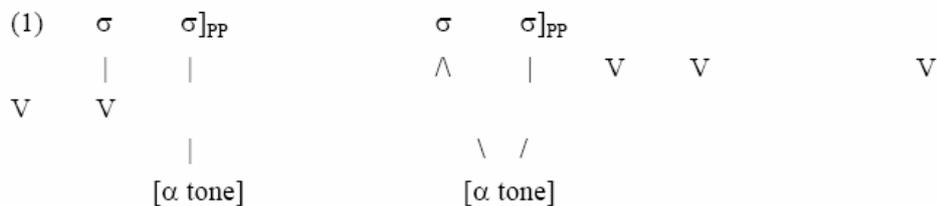
Humboldt-University of Berlin

bernd.pompino-marschall@rz.hu-berlin.de

1 Introduction

The aim of the study was to analyse a language with a simple tone system with respect to (i) intonational structure, (ii) vowel intrinsic fundamental frequency, and (iii) intonational marking of *in situ* focus.

Chichewa, a southern Bantu Language spoken in Malawi, has a simple tone system of H and L. On the surface contour tones HL and LH emerge through phrase penultimate lengthening and regressive tone spreading (cf. (1))



According to Kanerva (1990) *in situ* focus in Chichewa is marked by phrasing (cf. (2)); focus is marked by capitals)

- (2) (a) **anámén**ya **nyumbá** **ndí** **mwáálá** ‘S/he hit the house with a rock.’
 s/he hit house with rock
- (b) **(A-ná-mén**ya **nyumbá** **ndí** **mwáálá**).
 [neutral declarative]
- (c) **(A-ná-mén**ya **NYUÚMBÁ**) ! **(ndí** **mwáálá**).
 [Answers the question: *A-ná-ménya chiyáani ndi mwáálá?*
 ‘What did he hit with the rock?’]
- (d) **(A-ná-mén**ya **nyuúmbá**) **(NDÍ** **MWÁÁLÁ**). [no downstep]
 [Without downstep, emphasizes ‘rock’, but ‘house’
 (as it is postverbal and not coreferenced by an OM) is also new
 information]
- (e) **(A-NÁ-MÉENYA)** ! **(nyuúmbá)** ! **(ndí** **mwáálá**).
 [focus on hitting]

2 Method

Material

Five repetitions were read by a native speaker of Ntcheu Chichewa of sentences in neutral pronunciation and of different in situ focus conditions (focus position graphically marked or induced by question answer pairs) as well as single words. These were recorded and analysed with PRAAT.

Segmentation

The material was annotated with PRAAT at the word/pause level and the segmental level with respect to vowels.

Analyses

Besides durational measurements of words and pauses, the vowel segments were analysed with respect to duration and f_0 .

3 Results

Intonation of neutral declaratives

When normalized to utterance length, vowel fundamental frequency shows clear declination with

$$f_0 = 138.867 - 0.463 * \text{vowel position for unfocussed short high and}$$
$$f_0 = 110.517 - 0.239 * \text{vowel position for unfocussed short low vowels.}$$

Declination in focussed vowels

Focussed high vowels exhibit higher f_0 , nearly parallel in declination in comparison to their nonfocussed counterparts **Intrinsic f_0**

When measured fundamental frequency is corrected for this observed declination, there is a highly significant influence of categorical vowel height on f_0 .

Intonational focus marking

Besides phrasal focus marking, which results in long penultimate vowels (cf. above), there is clear intonational focus marking in utterances containing low and high tone syllables: *Focussed phrases are produced with higher fundamental frequency.*

In all low tone utterances no intonational focus marking can be found.

References

- Kanerva, J.M. (1990), *Focus and Phrasing in Chichewa Phonology*. New York: Garland.
- Downing, L.J., Mtenje A. & Pompino-Marschall, B. (2004), “*Prosody and information structure in Chichewa*“. *ZAS Papers in Linguistics* **37**, p. 167-186.
- Downing, L.J., Mtenje, A., Pompino-Marschall, B. (2005), “*Non-accentual Prosodic Cues to Focus in a Tone Language: the case of Ntcheu Chichewa*“. Paper read at the Between Stress and Tone Conference, Leiden University, 16-18 June 2005

Prosodic Edge Marking in Ewe

Stefanie Jannedy
(with support of Ines Fiedler) *

Humboldt University of Berlin

jannedy@ling.ohio-state.edu

1 Introduction

Tone languages such as Mandarin Chinese (Xu, 1999) and the Bantu language Chichewa (Downing, 2003) display tonal properties beyond the lexical level when some part of the utterance is in focus. Ameka (1992) describes the Gbe language Ewe to utilize morpho-syntactic means to signal focus without making any reference to prosody. The question arises whether there are typological commonalities with regard to prosody across different types of tone languages when signaling focus. Möhlig's (1971) makes reference to '*expressive prosodemes*' in Ewe which serve to emphasize a word or phrase or intensify the main meaning. We tested for these '*expressive prosodemes*' by comparing the phonetic realization of the high-toned /-é/ focus marker (FM) in subject focus condition and in ex-situ object focus utterances to see if the language also makes use of prosodic (durational or pitch) cues to highlight or structure information. While duration appears to mark phrasal boundaries after the focussed constituents, F0 measurements of controlled tonal patterns were inconclusive as of now (and thus are not discussed here).

* The people who advised and commented on the study are too numerous to mention by name, we thank them sincerely. Any errors are entirely our own.

2 Methods

The corpus consisted of simple SVO (out of the blue, subject focus and *in-situ* object focus) utterances, and OSV *ex-situ* object focus sentences, controlled for tonal co-occurrences on the lexical level as we were looking for prosodic effects of focus on the fundamental frequency (F0) and duration. Six tonal patterns were selected for further investigation of the following 3 conditions:

- | | | | | |
|----|---------------|----------|----------|-----------------|
| 1. | S | V | O | (36 utterances) |
| 2. | S + FM | V | O | (36 utterances) |
| 3. | O + FM | S | V | (39 utterances) |

The corpus was read by a single native speaker of the Anglo dialect of Ewe. Thus, from our descriptive phonetic study, for methodological reasons we cannot generalize over the entire language nor to all speakers: our findings pertain to this one speaker and serve as a starting point for further investigation of the morpho-phonological system of Ewe.

Phonemic segmentation was carried out according to standard measurement criteria: Double articulated consonants such as /gb/ were labeled as single phonemes. Release bursts were included with the stop, not the following vowel. Two adjacent vowels were segmented by analyzing changes in the formant trajectories or dips in the amplitude.

3 Results

We measured the durations of each phoneme in all sentences in the different focus contexts. In the argument focus conditions (2) and (3), we also measured the duration of the high-toned FM /-é/ at the right edge of the subject or *ex-situ* object (see focus conditions). In Ewe, there is ambiguity in the interpretation

between in-situ object focus without morpho-syntactic marking and *out-of-the-blue* utterances.

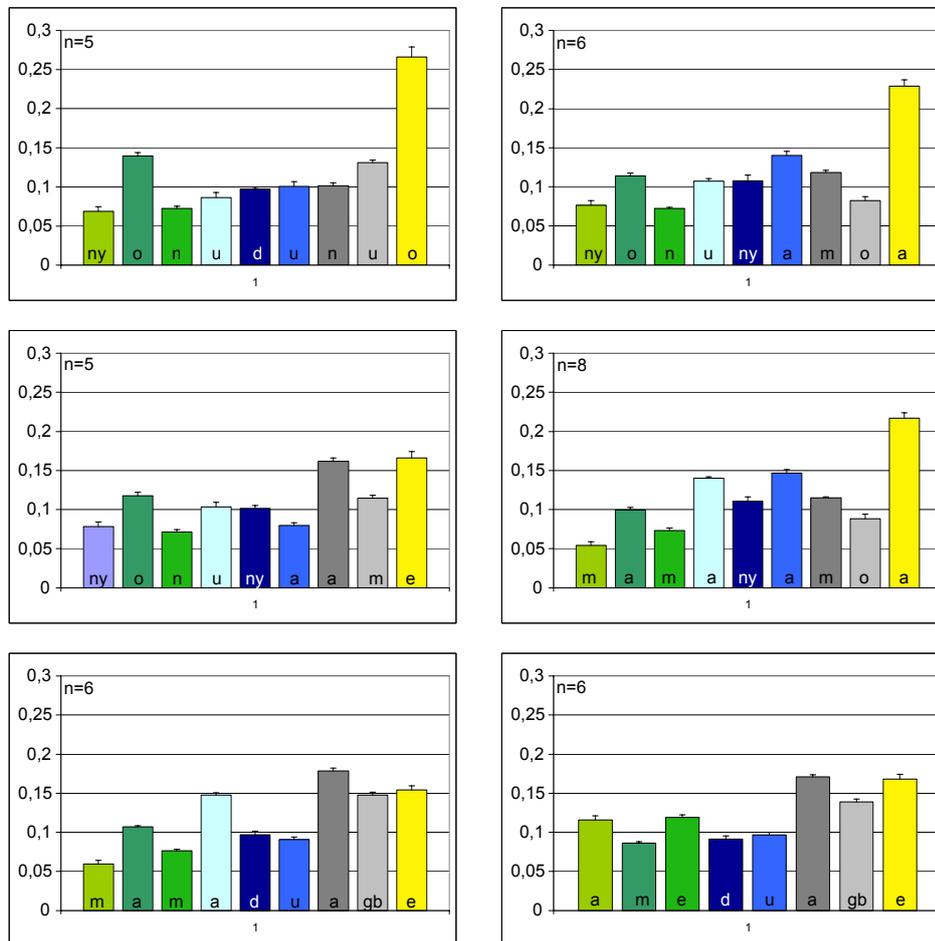


Fig 1: Mean duration (with standard error bars) of each phoneme (in seconds) for each tonal pattern

Regardless of these pragmatic differences, visual inspection of the mean phoneme duration across the six tonal contours (Fig. 1) indicates that only the final vowel is systematically elongated, suggesting that our Ewe speaker does final lengthening, a phonological process by which the right edge of a phrase is lengthened. Durational cues also play a significant role in structuring and/or highlighting the information in the S-Foc and O-ex-situ conditions: an ANOVA (linear mixed effects model with S+FM vs. O+FM as the fixed effect and the tonal contours as the random factor) shows a significant effect on the dependent variable ‘the duration’ of the focus marker ($p < .05$, $df=1$, $F=15.77$).

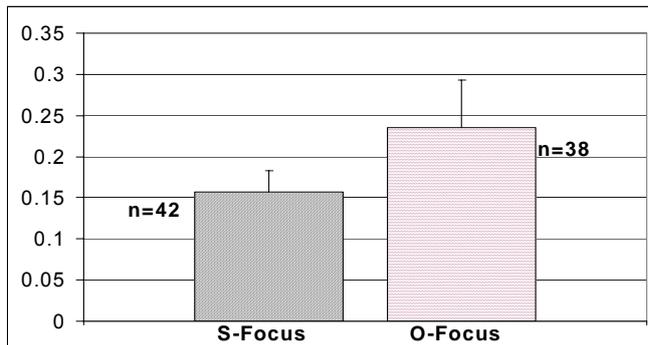


Fig 2: Mean duration (with standard deviation bars) of FM (in seconds) for S-Foc and O-ex-situ utterances.

This suggests a stronger phonological boundary between the ex-situ focused object and the remaining SV complex compared to the focused S and the VO.

4 Discussion

There is evidence for Möhlig's '*expressive prosodemes*' in the speech of our Ewe speaker: lengthening can have a perceptual effect of increased salience resulting in an interpretation of emphasis. Both S and O are set off prosodically when focussed, suggesting that the S could also be interpreted as *ex-situ*. This analysis in turn supports the *narrative hypothesis* (Fiedler & Schwarz), suggesting that historically, focus constructions in Ewe can be regarded as bi-clausal constructions, consisting of an NP and a narrative clause.

References

- Ameka, F. (1992) Focus Constructions in Ewe and Akan: A Comparative Perspective. In Collins, Ch. & Manfredi, V. (eds.) *Proceedings of the Kwa Comparative Syntax Workshop*, pp. 1-25. Cambridge.
- Downing, L. J. (2003) Stress, Tone and Focus in Chichewa and Xhosa. In Rose-Juliet Anyanwu, (edd) *Stress and Tone – the African Experience*. *Frankfurter Afrikanistische Blätter* 15, 59-81.

-
- Fiedler, I. & A. Schwarz. (to appear). Focus or Narrative Construction? In Aboh, E., Hartmann, K. & Zimmermann, M. (eds.) *Focus Strategies: Evidence from African Languages*, Berlin: de Gruyter.
- Möhlig, W. J. (1971) Zur Prosodologischen Struktur des Standard-Ewe. In Six, V., Cyffer, N., Wolff, E., Gerhardt, L. & H. Meyer-Bahlburg (eds.) *Afrikanische Sprachen und Kulturen: ein Querschnitt*. Deutsches Institut für Afrikaforschung, Hamburg.
- Xu, Y. (1999). Effects of Tone and Focus on the Formation and Alignment of F0 Contours. *Journal of Phonetics* 27: 55-105