

# FOCUS-TO-STRESS ALIGNMENT IN 4 TO 5-YEAR- OLD GERMAN-LEARNING CHILDREN

ANJA MÜLLER, BARBARA HÖHLE, MICHAELA  
SCHMITZ, AND JÜRGEN WEISSENBORN

## 1. Introduction

Children do not only have to learn how to express and to interpret the propositional content of a sentence, but also what is supposed to be common knowledge for the interlocutors and what is new information. Previous research has shown that from the beginning of multi-word utterances children seem to adequately use the prosodic, lexical and syntactic means of the target language to mark information structure, like for example focus stress, focus particles, and word order to indicate the focused element (e.g. Jannedy, 1997; Penner, Tracy & Weissenborn, 2000; Nederstigt, 2001). On the other hand, studies looking at comprehension performance suggest that children have difficulties, possibly beyond preschool age, interpreting a sentence containing focus particles like *only*, or with the interpretation of stress as a focus marker (e.g. Drozd, 2005; Drozd & van Loosbroek, 1998; Bergsma, 2002; Gualmini, Maciukaite & Crain, 2003; Paterson, Liversedge, Rowland & Filik, 2003; Hüttner, Drenhaus, van de Vijver & Weissenborn, 2003).

Why do 4-5 year and older children still have problems interpreting a given sentence in context correctly? One of the reasons for this may be that they have difficulties determining what is given and what is new, i.e. focused information with respect to the preceding linguistic context<sup>1</sup>. In order to investigate this hypothesis, we started from two widely accepted assumptions. First, that the focus of a particular sentence, as a rule, can be determined on the basis of the preceding context, and second, that there is a systematic relation between the informational focus of the sentence and its intonation, in the sense that the prominent accent (nuclear accent) has to be assigned to the rightmost element of the focused constituent, a principle called Focus-to-Stress-Alignment (FSA) by Nespor and Guasti (2002) following Jackendoff (1972). That contextual information and prosodic information, i.e. nuclear sentence accent are

essentially overlapping, as observed among others by Büring (to appear), is shown by the fact that when reading aloud a text, we automatically and rather consistently assign the appropriate focus intonation to it.

Thus, exploiting this redundancy, one way to find out whether a child is able to correctly determine the information structure of a given sentence in context, more specifically whether she is able to identify the new, i.e. the focused information, is to investigate whether she is able to realize the corresponding Focus-to-Stress Alignment.

A linguistic context with respect to which the semantic focus of the following sentence can be unambiguously determined is constituted by the question in a *wh*-question-answer pair<sup>2</sup>. That is, the focused constituent of the answer corresponds to the constituent in the scope of the *wh*-phrase in the question and thus has to be assigned the nuclear accent. This is illustrated by the following examples:

- (1) a. What happened? *John bought a CAR.*  
 b. What did John do? *John bought a CAR.*  
 c. What did John buy? *John bought a CAR.*  
 d. Who bought a car? *JOHN* bought a car.

The constituent underlined is focused, the large capitals mark the constituent carrying the nuclear accent. Thus in (1a) the focus consists of the answer as a whole, the so-called "broad focus", in (1b) of the VP, and in (1c) and (1d), of the object and the subject respectively, the so-called "narrow focus".

In a sentence with broad focus like (1a) which could also be uttered out of any context, the nuclear accent universally seems to fall on the rightmost element of the prosodic structure, whereas the syntactic constituent structure realizes the unmarked, canonical order, corresponding for example to *SVO* in languages like English, German, and Italian. We may consider the way focus-to-stress alignment (FSA) is realized in (1a) as the default, or unmarked case.

But languages vary in how they realize FSA in case of narrow subject focus. They may basically be differentiated with respect to whether they tend to maintain either the unmarked, canonical prosodic structure by choosing a word order that places the focused element at its right edge where nuclear accent is assigned according to the regular stress rule, as for example in Italian, or, alternatively, whether they assign the nuclear accent to the focused element in its syntactic position, thus abandoning the unmarked, canonical prosodic structure, as for example in English (e.g. Nespor & Guasti, 2002; Samek-Lodovici, 2005; Vallduví, 1993). This division basically corresponds to the traditional distinction between flexible/free and rigid/fixed word-order languages. Languages like German in which both word order and the position of

nuclear accent are relatively flexible, occupy an intermediate position. These different strategies are illustrated by the following examples from English, Italian, and German:

- (2) a. Who bought a car? *JOHN* bought a car. vs. \*bought a car *JOHN*.  
 b. Chi comprò una machina? (Una machina,) la comprò *GIOVANNI*.  
 c. Wer kaufte ein Auto? *PETER* kaufte ein Auto. or Ein Auto kaufte *PETER*.

Thus in English, the FSA is satisfied by stress movement, in Italian by movement of the subject from the canonical to the post-verbal, final position where nuclear accent is assigned according to the regular stress rule, whereas German has both options.

But on the other hand, German has less flexible word order in intransitive sentences than Italian as shown in (3) (e.g. Nespor & Guasti, 2002):

- (3) a. Chi è morto? È morto CALVINO  
 b. Wer ist gestorben \*ist gestorben HANS  
 (Who has died? has died Hans)

Summarizing so far we can say that the language-specific FSA pattern that we observe in the answers to narrow focus subject wh-questions make apparent which are the principles of FSA underlying these pattern, i.e. nuclear stress movement or focus movement. This also means that investigating the intonation pattern that German-learning children of age 4-5 assign to prosodically neuter test-sentences, displaying the different legitimate word-order pattern which are licensed by the preceding narrow focus subject and object wh-questions, should allow us to tell not only whether the children have correctly identified the focus of these sentences, i.e. what constitutes new information with respect to the preceding context, but also, whether they have acquired the principles determining FSA in German.

If they have, they should assign nuclear accent, i.e. the most prominent pitch accent to the focused constituent, and they should do so independently of whether the focused constituent occupies the sentence final position, where nuclear accent is canonically assigned by the regular stress rule, or whether it occupies a non-final sentence position, where nuclear accent can only be assigned through movement of the main prominence.

What would we predict if the children were not able to determine the focus structure of the answer-sentences? One possibility could be that they then analyze the sentence as unrelated to the preceding context, as an "out of the

blue” utterance, or like a sentence with broad focus answering the question “What happened?”. We have seen that these focus-neuter sentences universally realize main prominence on the rightmost constituent. This should then also be the predominant intonation pattern realized by the children. We would also predict that in this case, given that focus-neuter sentences, as a default, realize the canonical word order pattern of the given language, children should then possibly reorder non-canonical orders such that they conform to the canonical word order of German, i.e. SVO, when repeating the test sentence independently of whether the reordered element in the given context should be focused or not.

## **2. Method**

### **2.1 Participants**

Fifteen monolingual German children with a mean age of 4;09 years were tested (range: 4;01 years to 5;10 years).

The 7 girls and 8 boys attended different kindergartens in Potsdam. According to a questionnaire filled out by the parents none of the children had ever shown any indications of an impairment or delay in language acquisition.

Furthermore, a control group consisting of 14 female adult speakers was tested.

### **2.2 Design and Material**

All in all there were 32 simple target sentences, which were presented orally and which were to be repeated by the participants. All sentences consisted of a subject, a direct object and a verb. The sentences differed with respect to their constituent order and with respect to the focused constituent. Half of the sentences were syntactically canonical (subject-verb-object) and the other half was syntactically non-canonical (object-verb-subject). The subject was the focused constituent in half of the sentences while it was the object in the other half. Both factors (focus / word order) were completely crossed leading to the four experimental conditions shown in Table 1. 8 test sentences were constructed for each condition.

	Syntactically canonical	Syntactically non-canonical
Subject focus	8 items (FS)O <i>Eva</i> kauft die Gurke. <i>Eva</i> buys the cucumber. subject            object	8 items O(FS) Den Opa begrüßt <i>Peter</i> . The grandfather greets <i>Peter</i> . object                            subject
Object focus	8 items S(FO) Eva putzt <i>die Fenster</i> . Eva cleans <i>the windows</i> subject            object	8 items (FO)S <i>Die Treppe</i> fegt Peter. <i>The staircase</i> sweeps Peter object                            subject

**Table 1: Experimental conditions**

S=subject, O=object. F= focus

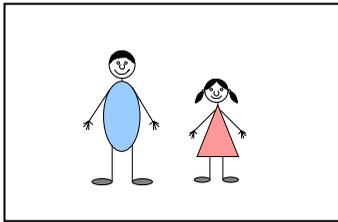
(FS)= subject should be focus marked

(FO)= object should be focus marked

For each of the 32 target sentences a short comic strip consisting of three pictures each was drawn. In addition, a short three-sentence story related to the pictures was created that provided the relevant contextual information. The story was followed by a question related to the last picture of the sequence. This question was followed by the answer to exactly this question (see Figure 1). This answer served as the target sentence for the imitation task.

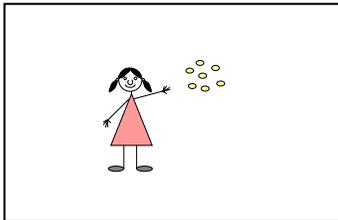
All verbal material for the experiments – except the target sentence – was recorded by a female native speaker of German. She was instructed to produce the sentences in a lively, child directed manner. To avoid any focus related prosodic information in the target sentences, these were spliced together from words recorded in isolation. For these means, the same female speaker that recorded the contexts and questions recorded all words necessary for the target sentences from a list in which these words had been ordered randomly. After recording, the single words were spliced out and put in the order of the target sentences. These sentences were again manipulated using a PRAAT script, so that the F0-value for each word of the sentence was set to 150 Hz.

All visual and acoustic material necessary for the experiment was transferred to a notebook that displayed the comic strips as well as the corresponding verbal stimuli.



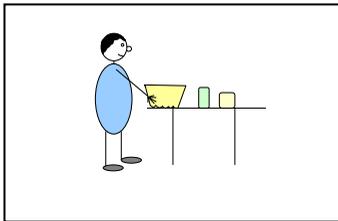
Morgen hat die Mama von Peter und Eva Geburtstag. Beide wollen deshalb für die Mama eine Überraschung vorbereiten.

Tomorrow is Peter's and Eva's mother's birthday. Therefore they want to surprise their mother.



Eva möchte Plätzchen backen.

Eva wants to bake cookies.



Was backt Peter?  
Peter backt Kuchen.

What does Peter bake?  
Peter bakes a cake.

**Figure 1: Example for the material**

The study was designed as an elicited imitation task. This kind of procedure seemed especially useful for our study out of the following reasons. First, elicited imitation does not reflect a parrot-like passive reproduction of the stimulus but clearly reflects the linguistic competence of the child by systematic changes between the stimulus and the child's output (Höhle et al, 2001; Lust, Flynn & Foley, 1996). This might be especially true when – as it is the case in our study – the stimulus sentence lacks any sentence intonation. Since it would be quite unnatural to produce a sentence without intonation the spontaneous addition of prosodic information can be expected.

The experiment took place in a separate room of the children's kindergarten. Each child was tested two times. The first time the child got to know the experimenter. Also during this first session two pre-tests were conducted to check whether the child understood wh-questions and whether the child was able to produce sentences with non-canonical word order. All children tested were able to perform both tasks. At the end of this first session a hand-puppet, a rabbit, was introduced to the child. The rabbit talked to the child and said that

next week he would bring a wonderful comic strip with him into the kindergarten for the child to see.

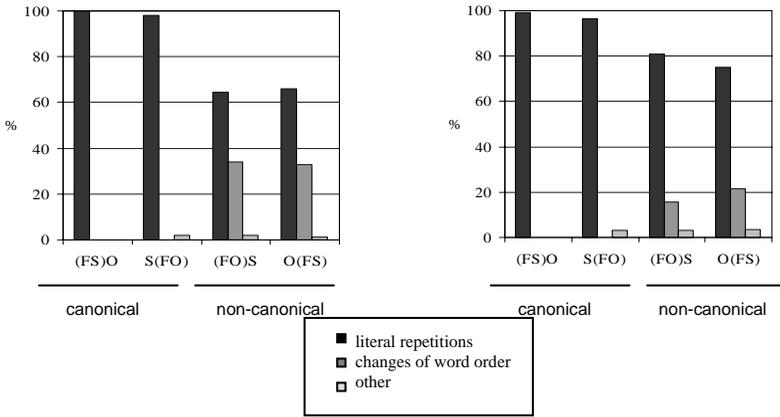
One week later the experimenter came a second time to the kindergarten. But this time the rabbit had a bandage on his head and of course the children asked what had happened. And then the rabbit told them that he was a bit ill at the moment, but that he nevertheless wanted to have a look at the comic strip together with the child. But because of the bandage the rabbit had some problems in understanding and therefore he asked the child to help him and repeat some things, because he might not understand every detail in the comic strip and in the story. If the child gave her consent, the rabbit and the child looked together at the comic strip, but after presenting the question and the target sentence the rabbit touched the child and said: "Oh, I didn't understand that," prompting the child to repeat the target sentence for the rabbit, using the prosody she thought to be appropriate.

For the adult control group there was only one test session because there were no pre-tests. In the test session we used exactly the same procedure for the control group as for the children. That means that the adults were also asked to help the rabbit and so they had to speak with him. All participants were tested individually and each test session lasted approximately 20-25 minutes. The verbal responses to the sentence to be imitated were recorded.

## 2.4 Data analysis and results

For the analysis all responses of the subjects were digitized for a further prosodic analysis using PRAAT as well as orthographically transcribed for an analysis of word order.

Both groups of subjects produced a high number of complete and literal repetitions of the target sentences (see Figures 2 and 3). The number of literal repetitions was statistically analyzed by a 2x2 factorial ANOVA (canonicity x subject/object focus) for the two groups of participants separately. The analysis of the children's data showed a significant main effect for the factor canonicity ( $F_{(1,14)} = 11.18$ ;  $p < 0.05$ ). No other effects reached significance. The same pattern emerged for the adults with the only significant effect for canonicity ( $F_{(1,13)} = 8.12$ ;  $p < 0.05$ ).

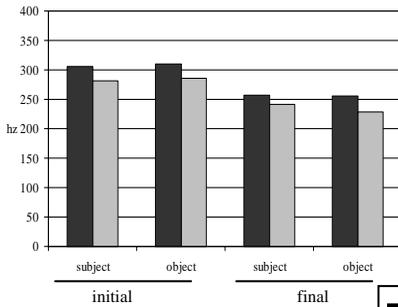


**Figure 2: Adults - Types of responses**

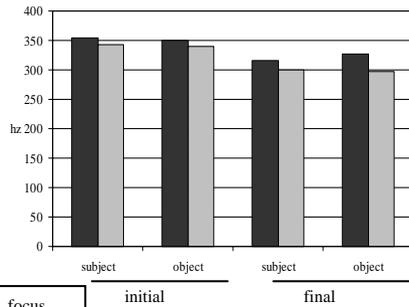
**Figure 3: Children - Types of responses**

Only the complete and literally imitated sentences were included in the prosodic analysis. We were able to conduct the prosodic analysis for all children but only for 10 of the adult controls. The other four adult participants systematically changed the word order in one of the non-canonical sentence conditions to a canonical word order so that for one sentence condition there is no analyzable material. We decided therefore to exclude these participants from the prosodic analysis.

Using PRAAT the pitch contour over the subjects and the objects of the imitations produced by the children and by the adults was determined. These pitch values were then averaged over the subjects and objects for the sentences belonging to the same experimental condition (see Figures 4 and 5).



**Figure 4: Adults – Average pitch measures**



**Figure 5: Children – Average pitch measures**

The statistical analysis of the data was conducted by a 2x2x2 (position x subject/object x focus) factorial ANOVA separately for the two groups of participants. The analysis for the children's data showed significant main effects for the factors position ( $F_{(1,14)} = 70.16$ ;  $p < 0.001$ ) and focus ( $F_{(1,14)} = 18.42$ ;  $p < 0.001$ ). All other effects or interactions failed to reach significance. The results of the adults' data revealed the same picture with significant main effects for position ( $F_{(1,9)} = 35.08$ ;  $p < 0.001$ ) and focus ( $F_{(1,9)} = 9.33$ ;  $p < 0.05$ ) and no further significant results.

Furthermore, we analyzed the number of changes of constituent order in our experimental conditions. Only responses that reversed the constituent order by keeping the subject/object functions of the arguments were considered in this analysis. As can be seen in Figures 2 and 3 neither the children nor the adults produced any changes of the constituent order in the conditions in which a canonical sentence had to be imitated. All changes of constituent order were observed in the non-canonical conditions changing a non-canonical target sentence to a canonical output. The number of these changes was numerically higher for the adults than for the children but there was no significant difference between the groups ( $t_{(df=27)} = 1.12$ ;  $p = 0.27$ ). The tendency for the children to produce more changes in constituent order in the subject focus condition than in the object focus condition also failed to reach significance ( $t_{(df=14)} = 1.13$ ;  $p = 0.27$ ).

### 3. Discussion

Summarizing our results we found that in the utterances of German 4-year-olds as well as German adults a focused element carries a higher pitch than an

unfocused element with the same syntactic function and the same position within an utterance. In addition we found a strong effect of position with respect to pitch: in both groups the initial constituent of the utterances always carries a higher pitch than the final one, irrespective of being focused or not. A second main finding of our study is the strong tendency for the production of sentences with canonical word order: the children as well as the adults show a tendency to produce canonical word order (SVO) irrespective of whether the subject or the object is being focused.

First of all, our results indicate that previous findings according to which English-learning children's productive behavior reflects an early mastery of the prosodic devices of focus marking (MacWhinney & Bates, 1978; Hornby & Hass, 1970; Jannedy, 1997) can be extended to children learning German. This suggests that the fact of the higher flexibility that German allows with respect to changes of word order or nuclear stress position does not necessarily make the task of acquiring the FSA rules harder. In addition, our experimental data suggest that children of the age we tested have no problems in using contextual information – at least in restricted contexts like question-answer pairs – for the identification of the focus of an utterance. This suggests that the problems children of the same age still show with the interpretation of sentences containing focus particles reported in the introduction (Gualmini et al., 2003) are not due to a general inability to infer information structure from given contexts.

In addition, our results suggest that children use linguistic means to express information structure – at least in the area under investigation, namely focus stress – adequately. This is worth noting since – as our pitch analysis also reveals – the focused constituent in a sentence must not carry the highest average pitch in the sentence. As our data have shown, the sentence initial constituent is systematically higher pitched than a final constituent. This fact is probably due to declination, i.e. the typical continuous decrease of the F0-contour within an intonational phrase due to physiological facts (e.g. Cohen & t'Hart, 1967). This means that the listener cannot simply rely on pitch peaks to identify the prosodic focus in a sentence but she must take declination into account to compute the most salient constituent in a sentence. Adult listeners seem to compensate for this declination easily: in a flat intonational contour a later appearing element is rated as being more stressed than an earlier one (Lieberman & Pierrehumbert, 1984). The child must have these compensatory processes available by a very early age otherwise it would be hard to figure out that focused elements in a sentence are prosodically highlighted. The fact that the children show the same effect for position and focus as the adults suggests that they have mastered the interplay of declination and focus stress in the planning and the motoric execution of the intonational pattern of a sentence.

This early mastery of prosodic means of focus marking might be supported by specific features of infant-directed speech. Fernald and Mazzie (1991) as well as Fischer and Tokura (1995) found a much more consistent placement of focused words on pitch peaks of the utterance in infant-directed speech as compared to adult-directed speech. In addition, in adult-directed speech prosodic prominence was conveyed by a variable composite of acoustic features, including pitch, duration and amplitude while pitch was the most prominent feature for focus marking in their infant-directed speech samples. This could suggest that infant-directed speech is characterized by features that make the signal more transparent with respect to form-function relations, which helps the infants to track the relevance of a single cue for specific functions.

For the moment, we only looked at pitch as one of the prosodic correlates that mark stress, because according to Pierrehumbert & Hirschberg (1990) the pitch accent marks the most prominent, i.e. in our case the focused lexical item. But further analysis should be extended to other cues that have been discussed as being relevant as acoustic correlates of prosodic prominence such as for example durational cues and intensity.

The second main result from our study concerns the strong tendency to produce canonical sentence structures that is obviously not influenced by the sentence's information structure. This result is remarkable since the children as well as the adults deviated in their reactions from the target sentence even though this was a clear violation of the directions given. The strong preference for the canonical word order is not only evidenced by the changes of the constituent order if the target sentences had a non-canonical structure but also by the fact that changes of constituent order did not in a single case occur when a canonical sentence was the target sentence.

This finding is remarkable given the fact that German – as described in the introduction – is a language in which both the constituent order as well as the placement of the nuclear stress are relatively flexible. Nevertheless, the results of our task suggest that both options are not equally acceptable or equally likely to be produced at least in the kind of experiment and for the type of sentences used. A follow-up experiment we did with 18 adult speakers suggests that it is not simply the task of reproducing the sentences or the experimental situation involving the "hearing impaired" rabbit that invoked the preference for canonical structures. In this follow-up, the experiment was run in the same fashion except that the subjects had to answer the question spontaneously and not to repeat the answer to the questions. In this task there were only three cases out of 576 answers that had a non-canonical word order. Overall, this suggests that, in the types of sentences we used, a more rigid word order seems to win over a rigid placement of the nuclear stress, putting German closer to languages like English than to languages like Italian. Further research should be extended

to other structures in German, especially on the ordering of dative and accusative objects within the middle field where a higher variability of word order might be expected. In addition, factors like animacy and definiteness are highly relevant for constituent ordering in German. These factors also might have contributed to the strong tendency for canonical sentences observed in our data.

### **Notes**

<sup>1</sup> For an operational definition of „given“ see Schwarzschild (1999).

<sup>2</sup> It is not excluded that a sentence in a given context may have more than one potential semantic focus as noticed by Bolinger (1972) cited in Büring (to appear).

### **Acknowledgements**

This study was supported by the German Research Foundation (DFG) as part of the SFB 632 “Information structure: the linguistic means for structuring utterances, sentences and texts” with a grant to Barbara Höhle and Jürgen Weissenborn. We thank our student assistants for their help in collecting and analyzing the data and Elizabeth Medvedovsky for assistance with the English text.

## References

- Bergsma, W. (2002) "Children's interpretations of Dutch sentences with the focus particle *alleen* ('only')," in I. Lasser, ed., *The Process of Language Acquisition. Proceedings of the GALA '99 Conference*, Peter Lang, Frankfurt (Main).
- Bolinger, D. (1972) "Accent is predictable if you're a mind reader," *Language* 48, 633-644.
- Büring, D. (to appear) "Intonation und Informationsstruktur," *Akten der IDS Jahrestagung 2005*, Mannheim.
- Cohen, A. and J. 't Hart. (1967) "On the anatomy of intonation," *Lingua* 19, 177-192.
- Drozdz, K. F. (to appear) "Children's Interpretation of Quantification," Special Issue of *Language Acquisition* 13 (3-4).
- Drozdz, K. F. and E. van Loosbroek (1998) "Dutch Children's Interpretation of Focus Particle Constructions," *Poster presented at the 23<sup>rd</sup> annual Boston University Conference on Language Development*, Boston, MA.
- Fernald, A. and C. Mazzie (1991) "Prosody and focus in speech to infants and adults," *Developmental Psychology* 27(2), 209-221.
- Fisher, C. and H. Tokura (1995) "The Given-New Contract in Speech to Infants," *Journal of Memory and Language* 34(3), 287-310.
- Gualmini, A., S. Maciukaite, and S. Crain (2003) "Children's Insensitivity to Contrastive Stress in Sentences with *ONLY*," in S. Arunachalan, E. Kaiser, and A. Williams, eds., *Proceedings of PLC 25* (Vol. PWPL 8.1).
- Höhle, B., J. Weissenborn, M. Schmitz, and A. Ischebeck (2001) "Discovering Word-Order Regularities: The Role of Prosodic Information for Early Parameter Setting," in J. Weissenborn and B. Höhle, eds., *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition* (Vol. 1), John Benjamins Publishing Company, Amsterdam.
- Hüttner, T., H. Drenhaus, R. van de Vijver, and J. Weissenborn (2003) "The acquisition of the German Focus Particle *auch* 'too': Comprehension does not always precede production", in A. Brugos, L. Micciulla, and C. E. Smith, eds., *BUCLD 28 Proceedings Supplement*, Boston, MA.
- Hornby, P. A. and W. A. Hass (1970) "Use of contrastive stress by preschool children," *Journal of Speech and Hearing Reserach* 13, 395-399.
- Jackendoff, R. (1972) *Semantic interpretation in generative grammar*, MIT Press, Cambridge, MA.

- Jannedy, S. (1997) "Acquisition of Narrow Focus Prosody," in *Language Acquisition, Knowledge Representation and Processing. Proceedings of the GALA '97 Conference*. Edinburgh, Edinburgh University Press.
- Lieberman, M. and J. Pierrehumbert (1984) "Intonational invariance under changes in pitch range and length," in M. Aronoff and R. Oehrle, eds., *Language and Sound Structure* MIT Press, Cambridge, MA.
- Lust, B., S. Flynn, and C. Foley (1996) "What Children Know about What They Say: Elicited Imitation as a Research Method for Assessing Children's Syntax," in D. McDaniel, C. McKee, and H. Smith Cairns, eds., *Methods for Assessing Children's Syntax*, MIT Press, Cambridge, MA.
- MacWhinney, B. and E. Bates (1978) "Sentential devices for conveying givenness and newness: A cross-cultural developmental study," *Journal of Verbal Learning and Verbal Behavior* 17, 539-558.
- Nederstigt, U. (2001) "The Acquisition of Additive "Focus Particles" in German," in A. H.-J. Do, L. Domínguez, and A. Johansen, eds., *Proceedings of the 25th annual Boston University Conference on Language Development*, Cascadilla Press, Somerville, MA.
- Nespor, M. and M. T. Guasti (2002) "Focus-stress alignment and its consequences for acquisition," *Lingue e Linguaggio* 1, 79-106.
- Paterson, K. B., S. P. Liversedge, C. Rowland, and R. Filik (2003) "Children's comprehension of sentences with focus particles," *Cognition* 89, 263-294.
- Penner, Z., R. Tracy, and J. Weissenborn (2000) "Where scrambling begins: Triggering object scrambling at the early stage in German and Bernese Swiss German," in C. Hamann and S. Powers, eds., *The Acquisition of Scrambling and Cliticization*, Kluwer, Dordrecht.
- Pierrehumbert, J. and J. Hirshberg (1990) "The Meaning of Intonation in the Interpretation of Discourse," in P. Cohen, J. Morgan, and M. Pollack, eds., *Intentions in Communication*, MIT Press, Cambridge, MA.
- Samek-Lodovici, V. (2005) "Prosody-syntax interaction in the expression of focus," *Natural Language & Linguistic Theory* 23, 687-755.
- Schwarzschild, R. (1999) "GIVENness, AvoidF and other Constraints on the Placement of Accent," *Natural Language Semantics* 7, 144-177.
- Vallduví, E. (1993) *The informational component*, Garland, New York