Prosodic Marking of Information Status in Picture Story Descriptions

Tabea Thies, Christine T. Röhr, Stefan Baumann, Martine Grice

IfL Phonetik, University of Cologne

{tabea.thies; christine.roehr; stefan.baumann; martine.grice}@uni-koeln.de

Abstract

In the present paper, we investigate the effect of information status on accent placement and accent types used in semi-spontaneous monologues elicited by picture story descriptions. Results confirm that newness is predominantly reflected in a more frequent use of nuclear pitch accents. However, counter to expectations, new referents are often produced with a falling accent, associated with low(er) prominence, while accessible and given referents are often produced with a rising accent, associated with higher prominence. Placed in the context of intonational phrasing, results reveal that new items are more likely to appear at larger phrase boundaries (intonation phrases (IPs)). Moreover, whereas in narratives of this type smaller (intermediate) phrases (ips) are typically marked by a H tone, IPs are typically marked with a L tone. This might explain the use of a falling pitch movement towards the low boundary, even for new items. This means, in our corpus there is a tendency for information status to determine phrasing, which determines the boundary tone type, which in turn affects the accent type.

Introduction

It is well-known that in intonation languages like German the information status of referents (or givenness) is marked prosodically. The concept of givenness assigns a particular information status (e.g. given or new) to individual discourse entities and the phrases that encode them with respect to the discourse context and/or the hearer's knowledge store (e.g. Lambrecht 1994). Newness in discourse has been shown to directly determine accent status (nuclear vs. non-nuclear) but also the level of prominence (defined here as the psycho-acoustic impression of 'standing out' in relation to neighbouring elements) of pitch accent types (Baumann, Röhr & Grice 2015): The lower the givenness of an item, the higher its prosodic prominence. However, such marking is less clear-cut in spontaneous speech than in read

speech (cf. Baumann & Riester 2013). Semi-spontaneous task-oriented speech, retaining a degree of naturalness whilst still allowing for some control, has been shown to be a promising source for research into information status and prosody. Röhr et al. (2016) reported a larger number of pitch accents and more prominent accents for marking newness in task-oriented *dialogues*.

The aim of the present paper is to collect further evidence for the effect of information status on accent placement and type from another kind of task-oriented speech, involving picture description *monologues*. For the present study, we generally hypothesize that the 'newer' (or less given) a discourse referent is the more it is made prosodically prominent by the speaker. More precisely, we expect to find intonational marking of information status, indicated by different distributions or probabilities of prosodic categories. The categories we are looking at are:

a) accent *placement* (assuming an increase in prominence (cf. Baumann 2014)): no accent < prenuclear accent < nuclear accent;

b) accent *type* in nuclear pitch accents (categorized according to GToBI (Grice, Baumann & Benzmüller 2005), assuming an increase in perceptual prominence (cf. Baumann & Röhr 2015)): no accent < falling accent (L*, H+L*, H+!H*) < rising accent (!H*, H*, L*+H, L+H*).

A discourse referent's information status is categorized according to the the *RefLex* annotation scheme (Riester & Baumann 2017). The categories used in our present study express an increase in the referential level of an item's newness from left to right: r-given < r-bridging < r-new (cf. Tab.2).

Method

Eight native German speakers told six different picture stories, each composed of eight single pictures. Each story contains two target items denoting discourse referents. One main target referent (**Item 1**) appears three times in a story while the other (**Item 2**) only appears

once (see Fig.1). The stories were arranged pairwise, i.e. six different target referents (see Tab.1) are embedded in pairs in two picture stories so that each target referent occurs as Item 1 in one story and as Item 2 in the other.
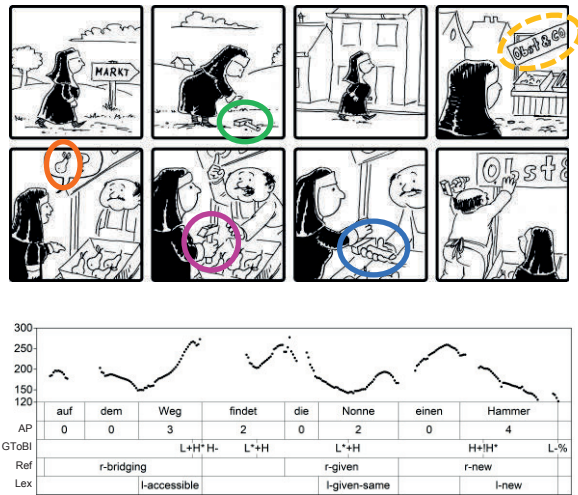

Figure 1. Sample of picture story 'hammer & pear' (top) with annotation in Praat of picture 2 (bottom).

The structure of the stories is as follows: Picture 1 = Filler, Picture 2 = **Item 1** (*new* discourse referent which is presented for the first time and is not derivable from the previous context), Picture 3 = Filler, Picture 4 = sets up a **Scenario**, Picture 5 = **Item 2** (*inferentially accessible* referent which has not been presented before but can be inferred via a bridging process from the contextual frame (see Tab.1)), Picture 6 = **Item 1** (*given-displaced* referent which is repeated after three intervening pictures with a change in topic), Picture 7 = **Item 1** (*given* referent which is a repetition of an antecedent in the immediately preceding context picture), Picture 8 = Filler.

Table 1. List of the six paired target items and their scenario condition.

Pair	Target referent	Scenario
1	pear (*Birne*)	fruit seller
	hammer (*Hammer*)	workshop
2	banana (*Banane*)	fruit shop
	needle (*Nadel*)	thimble
3	television (*Fernseher*)	living room
	screw (*Schraube*)	tool box

Experimental Setup

The experiment took place in a sound-attenuated room at the Cologne University phonetics lab and was composed of three parts a priming phase, a demonstration and training

phase and a test phase.

In the priming phase, the subjects were familiarized with the target items used in the picture stories. In the following main part of the experiment the whole picture story was first shown to the subjects by rapidly scrolling the eight pictures of a story on a computer screen. In a second step, all pictures were presented successively again and the subjects were instructed to tell the story in a natural, simple and coherent way without hesitation by describing the pictures. In a demonstration and training phase, subjects were familiarized with this procedure. In the final test phase, subjects told the six target picture stories in a pseudo-randomized order.

Subjects and Analysis

We recorded eight native speakers of Standard German (six female and two male). They were students of the University of Cologne aged between 23 and 27 years.

In a first step, we produced transcripts of the monologues. In a second step, we annotated the information status of all nouns according to the RefLex annotation scheme. At a referential level (Ref), labels were applied to DPs, PPs and APs. The information status of the words within these phrases were separately labelled at a lexical level (Lex). In this paper, we will focus on the referential level and its correlation with prosodic prominence (see Tab.2).

In a third step, we annotated the acoustic data in *Praat* (Boersma & Weenink 2017). We segmented every spoken word and analysed the prosodic realisation of all sentences at two different levels (see Fig.1). For accent placement (AP), we marked for each word whether it was realized without an accent (0), with a postnuclear accent (1), a prenuclear accent (2) or a nuclear accent (3 = in an intermediate phrase (ip), 4 = in an intonation phrase (IP)). At the tonal level (GToBI), we marked the positions of realized pitch accents and boundary tones and categorized their tonal configuration according to GToBI.

The results presented here are based on pooled GToBI and RefLex categories, even though we used the more fine-grained categories during the annotation process. For pitch accent types we distinguish between a falling (L*, H+L*, H+!H*) and a rising (H*, !H*, L*+H, L+H*) onglide (Grice et al. 2017). Two phrase types were annotated, the intermediate phrase (ip) and the intonation phrase (IP).

Boundary tones were pooled according to the beginning of their tonal movement towards the edge, i.e. we distinguish between low (ip: L-; IP: L-%, L-H%) and high (ip: H-, !H-; IP: H-%, H-^H%) boundary tones. RefLex categories were pooled according to the overview given in Table 2. For the entire scheme and detailed annotation guidelines consult Riester & Baumann (2017).

Table 2. Overview of simplified annotation tags of the referential level in the RefLex scheme.

Referential level (indicated by 'r-')	
r-given	coreferential anaphor that is present (immediately or displaced) in previous discourse context or contained in the text-external context
r-bridging	non-coreferential anaphor dependent on previously introduced scenario
r-new	globally unique or indefinite non-unique discourse-new (non-anaphoric) entity

Results

As an overall result, we found that the distribution of accent *placement* on all nouns as a function of their *referential* level of givenness confirms the expected incremental effect on an item's prosodic prominence (see Fig.2a).
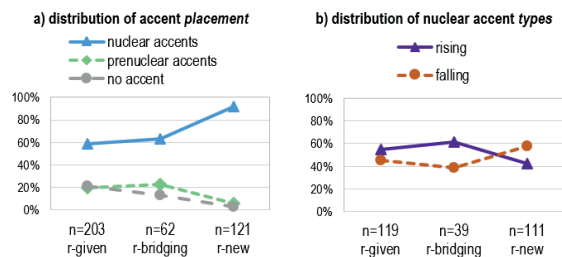

Figure 2. Relative distribution of a) accent placement and b) nuclear accent types on all nouns plotted against their referential level of givenness.

New referents more frequently receive nuclear accents than accessible (*r-bridging*) and given referents. Overall, we found a small proportion of cases with no accent or a prenuclear accent in this corpus. This is probably due to the short length of the utterances. Nevertheless, the possibility of no accent or prenuclear accent tendentially decreases with increasing discourse newness.

For the distribution of nuclear pitch accent *types* our expectations were not borne out (see Fig.2b). This is reflected by a reverse distribution of rising and falling accents, i.e. rising

accents (assumed to be most prominent) become less frequent, whereas falling accents become more frequent with increasing newness of a discourse referent.

If we combine these with aspects of intonational phrasing, we observe interesting patterns in our data. Typically, in German monologues (e.g. in story-telling) intermediate (i.e. smaller) phrases are marked by a H tone, indicating continuity, while the larger intonation phrases are marked with a L tone, expressing finality within a domain. Interestingly, new items typically occur at the end of intonation phrases, i.e. before a L tone. Our results confirm this: New items (in nuclear position) more often appear at larger phrase boundaries (see Fig.3) that are more often marked by a L tone, in particular when the nuclear item is a new referent (see Fig.4). Hence, in this corpus new items are often produced with a falling pitch movement, reflecting a following low boundary.
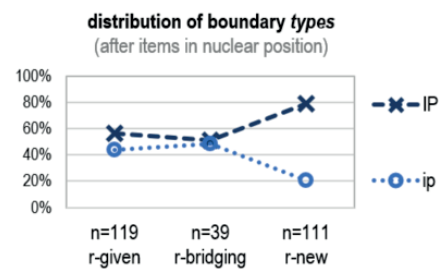

Figure 3. Relative distribution of intermediate (ip) and intonation (IP) phrases after nouns plotted against their referential level of givenness.
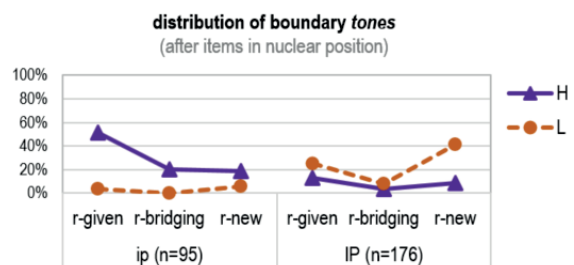

Figure 4. Relative distribution of high (H) and low (L) boundary tones of intermediate (ip) and intonation (IP) phrases after nouns plotted against their referential level of givenness.

Conclusions

We investigated (semi-)spontaneous speech in the form of task-oriented monologues which provide clear evidence for an interrelation between levels of referential givenness and their prosodic realisation. The data confirm that accent placement is a decisive prosodic marker of information status in that the distribution of accents shows a stepwise increase in prosodic

prominence from given and accessible to new referents. However, unexpectedly, our data reveal that new information is marked with nuclear accent types that are associated with lower perceptual prominence whereas accessible and given items are marked with nuclear accent types that are associated with higher perceptual prominence. This result is probably due to a pattern of intonational phrasing that is common in the text type investigated. Results confirm that new items are more likely to appear at larger phrase boundaries (IPs), which are typically marked by a L tone.

The end of the IP being a prominent position is in line with studies on working memory showing that items at the end of units are better recalled than items in other positions, making the final position prominent (*recency effect*: Jahnke 1965 inter alia). Furthermore, the choice of a falling pitch accent may be due to the polarity of tones (Hyman 2007), i.e. a pitch accent containing a H tone or falling tonal movement before a L edge tone, so that the low tone is approached from a higher pitch. Although in intonational phonology it is assumed that all pitch accents are freely combinable with all boundary tones, there are common collocations of tones that tend towards polarity, especially at boundaries (H tones before L, L tones before H, cf. Dainora 2001).

Thus, our data support the assumption that prominence is lent (to some extent) by the position in the phrase. Although we have found elsewhere that information status determines the pitch accent type directly, in this corpus information status determines phrasing which in turn determines the boundary tone type, and the boundary tone type in turn affects whether an accent is rising or falling. These results are interesting for Jun's (2006) typology of head marking and edge marking languages. Although German is a prototypical example of the former, it can also make use of edge marking for lending prominence.

Acknowledgements

References

Baumann, S. (2014). The Importance of Tonal Cues for Untrained Listeners in Judging Prominence. *Proceedings of the 10th ISSP*, 21-24.

Baumann, S., & Riester, A. (2013). Coreference, Lexical Givenness and Prosody in German. In J. Hartmann, J. Radó & S. Winkler (Eds.), [Special Issue "Information Structure Triggers"] *Lingua, 36*, 16-37.

Baumann, S., & Röhr, C. T. (2015). The Perceptual Prominence of Pitch Accent Types in German. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th ICPhS*, paper number 298, (pp. 1-5). Glasgow, UK: the University of Glasgow.

Baumann, S., Röhr, C. T., & Grice, M. (2015). Prosodische (De-)Kodierung des Informationsstatus im Deutschen. *Zeitschrift für Sprachwissenschaft, 34*(1), 1-42.

Boersma, P., & Weenink, D. (2017). Praat: Doing Phonetics by Computer [Computer program]. Version 6.0.28, retrieved 23 March 2017 from http://www.praat.org/

Dainora, A. (2001). *An Empirically Based Probabilistic Model of Intonation in American English.* University of Chicago, Department of Linguistics: PhD dissertation.

Grice, M., Baumann, S., & Benzmüller, R. (2005). German Intonation in Autosegmental-Metrical Phonology. In S. Jun (Ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing* (pp. 55-83). Oxford: OUP.

Grice, M., Ritter, S., Niemann, H., & Roettger, T. B. (2017). Integrating the discreteness and continuity of intonational categories. *Journal of Phonetics, 64*, 90-107.

Hyman, L. (2007). Universals of Tone Rules: 30 Years Later. In C. Gussenhoven & T. Riad (Eds.), *Tones and Tunes: Studies in Word and Sentence Prosody* (pp. 1- 34). Berlin: Mouton de Gruyter.

Jahnke, J. C. (1965). Primacy and Recency Effects in Serial-Position Curves of Immediate Recall. *Journal of Experimental Psychology, 70*(1), 130-132.

Jun, S.-A. (2006). Prosodic Typology: by Prominence Type, Word Prosody, and Macro-Rhythm. In S.-A, Jun (Ed.), *Prosodic Typology II: The Phonology of Intonation and Phrasing* (vol. 1) (pp. 520-540). Oxford: Oxford University Press.

Lambrecht, K. (1994). *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

Riester, A., & Baumann, S. (2017). The RefLex Scheme – Annotation Guidelines. In *Working Papers of the SFB 732, SinSpeC* (vol. 14) (pp. 1-31). University of Stuttgart.

Röhr, C. T., Thies, T., Baumann, S., & Grice, M. (2016). Prosodic Marking of Information Status in Task-Oriented Dialogues. In *Proceedings of the 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum* (pp. 168-171). München, Germany.