

Prosodic Marking of Information Status in German

Christine Röhr & Stefan Baumann

IfL Phonetik, University of Cologne, Germany

{christine.roehr; stefan.baumann}@uni-koeln.de

Abstract

In a production experiment on read German, we investigate the prosodic marking of discourse referents reflecting different types of information status. Acoustic and phonological analyses reveal an increase in the number of pitch accents as well as higher and later accentual peaks from textually given through textually accessible and inferentially accessible to new referents. Due to the increasing number of produced accents, segmental durations also increase from given to new information. Furthermore, specific accent types lead to different segmental durations. The differences in the prosodic marking of the two types of accessible information suggest a difference in cognitive activation between them, supporting the idea of an activation continuum of discourse referents.

Index Terms: prosody, information status, degree of givenness, cognitive activation, pitch accent, effort code

1. Introduction

The dimension of ‘given’ versus ‘new’ information is a central part in the investigation of information structure. Nevertheless, the various approaches to givenness in the literature differ with respect to the level this notion applies to (see [17] for an overview). In the present paper, we adopt Chafe’s [4], [5] cognitive view of givenness defined as the degree of activation of a referent or proposition assumed by the speaker to be in the listener’s consciousness at the time of utterance. Following Chafe, we postulate three different types of information status corresponding to three steps on a potentially continuous scale of cognitive activation: if a referent is already active in the listener’s consciousness at the time of the utterance, it is *given*; if a referent becomes activated from a previously semi-active state, it is *accessible*; and if a referent becomes activated from a previously inactive state, it is *new* (see Fig.1).

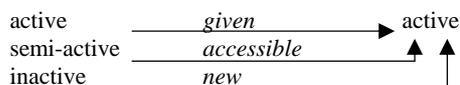


Figure 1: Chafe’s [5: 73] model of givenness degrees or types of information status.

There are at least three different origins the givenness of a referent may be derived from (see [15]). First, a referent may be recoverable because it is stored in the memory of speaker and listener, either as a unique referent which is part of a ‘public’ knowledge of the world (e.g. *the sun*) or as a more idiosyncratic piece of shared knowledge between the interlocutors (e.g. *John*). Second, a referent can be present in the text-external world, i.e. visible or otherwise salient in the speech setting. Third, a referent can be present in the text-internal world due to previous mention. This mention may be explicit, e.g. in a repetition of the same referent, or implicit, e.g. if the referent in question is part of a scenario (e.g. *courtroom – judge*) or stands in a hyponymy or meronymy relation to its antecedent (e.g. *dog – animal*). In the case of

explicit (co-)reference, we distinguish between immediately evoked items which we will call ‘textually given’, and items whose previous mention is non-immediate or ‘displaced’ (see [19]), referred to here as ‘textually accessible’. Implicit reference, which does not usually mean coreference, involves cognitive bridging [6] between an antecedent and an anaphor and will be subsumed under the term ‘inferentially accessible’ information.

In terms of the prosodic marking of discourse referents, it is commonly assumed for West Germanic languages that new referents are marked by pitch accents whereas given referents get deaccented (see e.g. [8]). The prosody of accessible referents, however, is a matter of some debate. Chafe [5], e.g., postulates that accessible information is marked – like new information – by accented noun phrases, while Lambrecht [15] suggests that accessible referents are either accented or deaccented. However, several studies have shown by now that a simple dichotomy of accentuation versus deaccentuation is inappropriate for an account of accessible information, and for an account of information status (or degrees of givenness) in general. Pierrehumbert & Hirschberg [16] for American English and Kohler [14] for German proved that the *accent type* or, respectively, the *tonal configuration*, are important cues for encoding a referent’s information status as well as higher-level semantic-pragmatic relations. In particular, Pierrehumbert & Hirschberg’s study suggests a ternary distinction between high accents for new, low accents for accessible and no accents for given referents (applying to the ‘hearer’s mutual beliefs’), while Kohler’s perception experiments indicate an interrelation between medial/late peaks and some kind of new information on the one hand and between early peaks and ‘established’ (interpreted here as accessible) information on the other.

In a recent perception experiment on German, Baumann [1] (see also [2]) could show that accessible information cannot be treated as a uniform category and that different types of more or less activated information demand different accent types as linguistic markers. In fact, there is evidence that a range of accent types (including deaccentuation) can be mapped onto the gradient scale of activation degrees, with the pitch height on the accented syllable being the determining factor. Such a mapping suggests a somewhat iconic use of pitch height, which is compatible with Gussenhoven’s [13] *Effort Code*: the higher the pitch on a lexically stressed syllable – and, in turn, the higher its prominence – the newer (or more newsworthy) the discourse referent.

In order to test this basic hypothesis in production data, we conduct a reading experiment displaying new, inferentially and textually accessible as well as textually given target referents. In addition to examining the types of pitch accent used, we investigate the alignment of the F0 peaks and valleys with the lexically stressed syllable of the target words, and the duration of these syllables. We hypothesize that speakers make use of less, and less pronounced, prominence-lending cues as the degree of cognitive activation increases. That is, new referents are expected to be marked by more and higher accents with

later peaks and longer segmental durations than accessible referents and, in turn, given referents. Within the group of accessible referents, we expect more prominence-lending cues on inferentially than on textually accessible items, since bridging inferences probably require more activation cost than the repetition of a, however displaced, referent (see [7] for a discussion).

2. Method

2.1. Reading material

The test material consists of ten different target words denoting discourse referents, each of them embedded in four target sentences in three different contexts (see Table 1).

Table 1. Example reading material for the target word Banane ('banana').

target word = Ba.na.ne [ba'na:nə] banana
<p>CONTEXT 1: (a) new (b) textually accessible</p> <p>„Was hätten Sie gerne?“. (a) „Ich nehme die Banane mit“, antwortet Thomas dem Obsthändler. Normalerweise ernährt er sich sehr ungesund und isst zwischendurch ständig Süßigkeiten. Außerdem treibt er fast nie Sport und wenn doch, dann am liebsten Minigolf. (b) Er steckt sich die Banane ein. Lecker sieht die Banane aus. Vielleicht wird er demnächst öfter welche kaufen.</p> <p>“What would you like?“ (a) “I’ll take the banana (along)“, says Thomas to the fruit merchant. He usually eats very unhealthily and he is always eating sweets between meals. He hardly ever plays sport, and if he does he prefers mini golf. (b) He pockets the banana. The banana looks delicious. Maybe he’ll buy them more often in future.</p>
<p>CONTEXT 2: (c) scenario (inferentially accessible)</p> <p>Thomas darf heute im Zoo seinen Lieblingsaffen füttern. Voller Vorfreude wird er sich gleich auf den Weg zu ihm machen. (c) Er steckt sich die Banane ein. Vorhin war er dafür extra noch auf dem Markt beim Obsthändler.</p> <p>Today Thomas is allowed to feed his favourite monkey in the zoo. With great anticipation he’s about to set off (for the zoo). (c) He pockets the banana. He’s just been to the green grocer’s at the market especially to get one.</p>
<p>CONTEXT 3: (d) textually given</p> <p>Thomas hat gerade auf dem Markt eine Banane gekauft. (d) Er steckt sich die Banane ein. In Zukunft möchte er sich viel gesünder ernähren.</p> <p>Thomas has just bought a banana at the market. (d) He pockets the banana. In the future he wants to eat much more healthily.</p>

The target words are bi- and tri-syllabic nouns (*Ballade* ‘ballad’, *Banane* ‘banana’, *Dame* ‘lady’, *Lawine* ‘avalanche’, *Rosine* ‘raisin’) and proper names (*Janina*, *Nina*, (*Dr.*) *Bahber/Bieber*, *Romana*) in feminine gender, always with stress on the penultimate syllable. In order to ensure segmental comparability, the target words are strictly composed of voiced sounds in open syllables, with the target syllable including one of the long vowels /i:/ or /a:/.

The structure of the test sentences is simple and kept constant in all contexts: Each target sentence starts with a pronominal subject followed by the finite part of the separable verb and the target word encoded as a definite direct object, and ends with the verbal particle (i.e. the prefix of the separable verb).

In target sentence (a), the target word is mentioned for the first time and is not derivable from the previous sentence. Thus, it can be assigned the information status *new*. After two or three intervening context sentences with a change in topic,

the target word is repeated in target sentence (b). Due to this displacement from the centre of attention, the target word is no longer fully activated and can thus be classified as *textually accessible* (cf. Centering Theory, e.g. [11]). The second context sets up a scenario, from which the referent in target sentence (c) is *inferentially accessible*. That is, the referent has not been mentioned before but can be inferred from the contextual frame (in Table 1: the banana is inferrable from the zoo-and-monkey context). Part of the third context is target sentence (d), which immediately follows a context sentence that includes the test word. This means that the target word is already fully activated and thus *textually given*.

As far as possible, we controlled the focus structure of the test sentences in order to keep its influence on the prosodic marking of the target words to a minimum. In target sentences (a), (b) and (c), the test words are part of a broad focus domain. Only in target sentence (d), the target referent is part of the background due to its mention in the immediately preceding context sentence.

2.2. Speakers and recordings

Nine native speakers of German (six female, three male), aged between 22 and 31, took part in the experiment. All of them originated from the area around Cologne and Düsseldorf. Before the acoustic recordings, each subject was asked to read through the material thoroughly in order to guarantee full comprehension. After that, the subjects’ task was to read out the texts in a contextually appropriate manner. The contexts were presented on separate file cards in pseudo-randomised order. They were repeated three times by each subject, adding up to 120 target sentences per speaker for analysis, i.e. 1080 tokens in total.

2.3. Analysis

The acoustic data were annotated manually using the EMU software [3]. At four levels, we marked the beginning and the end of the target word, the foot starting with the lexically stressed syllable, the stressed syllable itself and the vowel included in it.

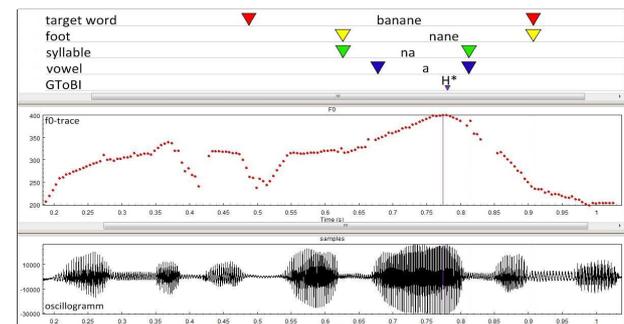


Figure 2. Target sentence *Ich nehme die Banane mit* ('I'll take the banana (along)'); new, speaker S08) with labels, F0 trace and oscillogram in EMU.

At the tonal level, we annotated the F0 minima and maxima making up pitch accents and categorised them according to GToBI [9], [10]. The structure of the target sentences, with the argument in non-final position, allows the nuclear accent either to fall on the target word or on the sentence-final verbal particle. In the latter case the target word is either deaccented (marked by '0') or receives a prenuclear accent (e.g. in *Er steckt sich die Banane EIN.*), indicated by an additional 'PN' symbol. The alignment of the F0 target with the stressed

syllable can be measured as the distance between the beginning of the syllable and the tonal label in relation to the duration of the whole syllable. An example screen shot with the described annotation levels is given in Fig.2.

3. Results and Discussion

3.1. Accent types

As an overall result, the distribution of accent categories (including prenuclear accents and deaccentuation) proves to depend significantly on the referents' information status (chi square: $p < 0.001$), as shown in Fig.3.

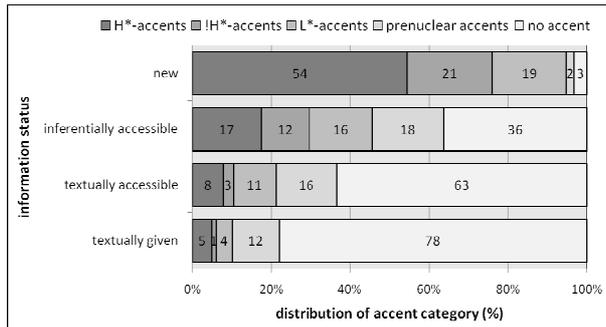


Figure 3: Relative distribution of nuclear accent categories (comprising accent types with the same starred tone), prenuclear accents and deaccentuation on all target words per information status; all speakers pooled.

Results show that, on the whole, new information is accented (97% of all referents) and textually given information is deaccented (i.e. it receives no accent, 78%). Accessible information takes an intermediate position. Interestingly, however, there is a clear difference between inferentially accessible referents, two thirds of which receive an accent, and textually accessible referents, two thirds of which are deaccented. Generally, we observe an increase in the number of pitch accents from textually given to textually accessible to inferentially accessible to new discourse referents.

Looking at accented material only, we find that new and accessible information is primarily marked by nuclear pitch accents. Given referents, if accented at all, are marked equally often by nuclear (10%) and prenuclear (12%) accents. Prenuclear accents on new referents are very rare (2%), which is an expected result for broad focus sentences, in which the argument usually receives the nuclear accent, not the predicate (at least in West Germanic languages, see e.g. [12], [18]). Although the two types of accessible referents occur in broad focus as well, they are to some extent marked by prenuclear accents (inferentially accessible: 18%, textually accessible: 16%). In fact, the relative preference for prenuclear accents compared with nuclear accents increases from new through inferentially and textually accessible to given referents.

Furthermore, the proportion of all three nuclear accent categories (H*, !H*, L*) increases from given to new information. This increase is particularly clear for H* accents which are most commonly used to mark new information (54%). For accessible and given information, lower accent categories become more important, with a relative tendency towards L* accents as opposed to downstepped H* accents as the referent's degree of activation increases (see Fig.3). Table 2 provides a closer look at the preferences for specific pitch accent types per information status used in the experiment.

Table 2. Relative distribution of accent types, including prenuclear accents (PN) and deaccentuation (0), on all target words per information status; all speakers pooled.

		information status (all speakers & all target referents): tonal marking									
preference for deaccentuation of target referent	low	new	accents	H* >	L+H* >	!H* >	H+!H* >	H+IH* >	L* >	0 >	PN
		% of 270	40,4	14,1	12,6	11,1	8,9	7,8	3,3	1,9	
high	inferentially accessible	accents	0 >	PN >	H* >	H+L* >	H+IH* >	!H* >	L* >	L+H*	
		% of 270	36,3	18,1	16,7	11,5	7,0	5,2	4,4	0,7	
high	textually accessible	accents	0 >	PN >	H* >	L* >	H+L* >	H+IH* >	!H* >	L+H*	
		% of 270	63,7	15,2	7,4	5,9	4,8	1,5	1,1	0,4	
high	textually given	accents	0 >	PN >	H* >	H+L* >	L* >	H+IH* >	L+H* =	!H*	
		% of 269	78,1	11,9	4,5	2,2	1,9	0,7	0,4	0,4	

As in the analysis of the comprised pitch accent categories, the distribution of single pitch accent types shows significant differences between the four types of information status (chi square: $p < 0.001$). The most prominent pitch accent types L+H* and H* are preferably used for marking new information. However, Table 2 shows that H* is the most frequent single type of accent on accessible and given referents as well, if they receive a nuclear pitch accent (the most frequent markers are deaccentuation and prenuclear accents, very often realised as L*). Thus, H* qualifies as the default nuclear accent type. Within the information status 'inferentially accessible', however, the two types of early peak accents (H+L* and H+!H*) form a larger group than H* accents (18.5% vs. 16.7%). In comparison, for textually accessible referents the two L* accent types (L* and H+L*) taken together are more frequent than H* accents (10.7% vs. 7.4%; see also Fig.3).

This distribution of accent types supports Kohler's [14] proposal that medial (H*) and late peaks (L+H*) rather mark new information while early peaks (H+!H*, H+L*) mark established information, which is comparable to our notion of accessibility (and givenness). In terms of alignment differences within the same accent type, we only find significantly later H* accents marking new versus textually accessible referents (ANOVA: $F(3, 186) = 8.197, p < 0.001$).

In sum, the variation in frequency, position and type of accentuation with changes in the referents' information status clearly confirms our hypothesis: the 'newer' the referent, the more likely its marking by a high (and late-peak) nuclear pitch accent. As the degree of activation of the target referent increases, subjects are more likely to use lower (and early-peak) and less prominent accents (if at all). The semantic and structural differences between the test words had no significant effect on their prosodic marking.

However, the results are to some extent speaker-specific. While new information generally gets accented, some subjects hardly differentiate between the prosodic marking of accessible and given information. There seem to be (at least) two different types of speaker: those speakers who usually deaccent accessible and given information generally prefer low pitch accents (for marking new information), while those speakers who often place an accent on accessible and given referents use high pitch accents by default.

3.2. Segmental durations

Results reveal a significant effect of information status on segmental durations, e.g. of the lexically stressed syllable of the target word (ANOVA: $F(3, 1080) = 20.214, p < 0.001$).¹ We find a stepwise increase in the duration of the target syllable from given through textually and inferentially

¹ We will restrict ourselves to the lexically stressed syllable here; analogous results are found for the other segmental domains investigated, the foot and the stressed vowel.

accessible to new referents. However, post hoc tests (Tukey-HSD) only show significant steps between textually given, inferentially accessible and new target words. The duration of textually accessible referents only significantly differs from new referents. Nevertheless, in terms of syllable durations, our hypothesis is generally confirmed.

Within the same accent type, however, no significant durational differences due to variation of information status can be found. That is, the change in duration of the stressed syllable depends on – or is a concomitant of – the speaker's choice of accent (see Fig.4).

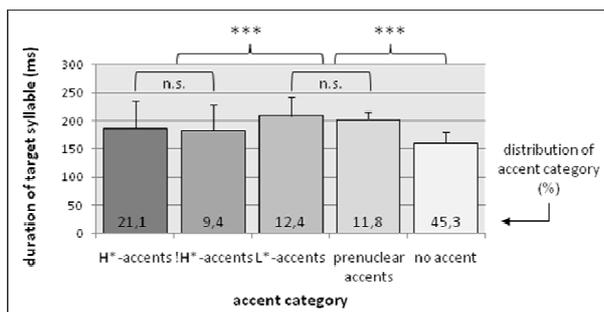


Figure 4: Durations of target referent's stressed syllable per accent category; all speakers pooled.

This means that different accent categories are characterised by different durations (ANOVA: $F(4, 1080) = 79.766$, $p < 0.001$). Surprisingly, however, post hoc tests (Tukey-HSD) show that syllables bearing L* and prenuclear accents (which are also often low) are significantly longer than syllables carrying H* and !H* accents, although the latter two were found to be the most frequent markers of new referents, which are generally longer than accessible and given referents (shortest syllable durations in deaccented target words). In fact, this result confirms the claim that for low accents (nuclear and prenuclear) duration is a decisive prominence-lending feature compensating for a lack of tonal movement (see e.g. [9: 278]).

4. Conclusions

Generally, our hypotheses are confirmed. The newer, or less activated, a referent, the more likely it is to be marked by a nuclear pitch accent in read German. Conversely, the higher the degree of a referent's activation, the higher is the preference for deaccentuation. Prenuclear accents are only used if the referent is already accessible or given in the discourse. As for the types of accent used, new information is found to be primarily marked by high and relatively late peaks while in accessible and given information the relative proportion of lower and early peak accents increases. Since more pitch accents are produced on less activated referents, a decreasing degree of givenness is reflected by longer segmental durations. In fact, the duration of the target word's stressed syllable partly depends on the type of accent used, with – somewhat surprisingly – syllables carrying low accents being longer than syllables carrying high accents. Thus, segmental durations only indirectly reflect a constituent's degree of activation.

As a main result of the study, the stepwise change in the relation between accentuation and deaccentuation among the four types of information status investigated suggests a difference in cognitive activation between the two types of accessible information. As expected, we find more

prominence-lending cues on inferentially than on textually accessible items, which seems to confirm the hypothesis that a bridging inference between an anaphor and its antecedent involves more activation cost than the explicit repetition of a referent. Furthermore, the observed differences in the prosodic marking of discourse referents within the same information status (here: accessible information) indirectly support our basic assumption that the system of cognitive activation of information is a continuum.

Our results show that the actual prosodic marking of a referent ultimately depends on speaker-specific preferences. Nevertheless, the idea of activation degrees, and in particular their marking by corresponding degrees of prominence, indicated by pitch height and peak alignment, is highly compatible with Gussenhoven's [13] *Effort Code*: higher and later peaks indicate less active referents.

5. References

- [1] Baumann, Stefan. 2006. *The Intonation of Givenness - Evidence from German*. Linguistische Arbeiten 508. Tübingen: Niemeyer.
- [2] Baumann, Stefan & Martine Grice. 2006. The Intonation of Accessibility. *Journal of Pragmatics* 38 (10), 1636-1657.
- [3] Cassidy, Steve & Jonathan Harrington. 2001. Multi-Level Annotation in the EMU Speech Database Management System. *Speech Communication* 33 (1-2), 61-78.
- [4] Chafe, Wallace. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. In: Charles Li (ed.), *Subject and Topic*. New York: Academic Press. 25-56.
- [5] Chafe, Wallace. 1994. *Discourse, Consciousness, and Time*. Chicago/London: University of Chicago Press.
- [6] Clark, Herbert H. 1977. Bridging. In: P.N. Johnson-Laird & P.C. Wason (eds.), *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press. 411-420.
- [7] Clark, Herbert H. & Susan E. Haviland. 1977. Comprehension and the Given-New Contract. In: Roy Freedle (ed.), *Discourse Production and Comprehension*. New Jersey: Ablex. 1-40.
- [8] Cruttenden, Alan. 2006. The De-accenting of Given Information: a Cognitive Universal? In: Bernini, G. & M. L. Schwartz (eds.), *Pragmatic Organization of Discourse in the Language of Europe*. The Hague: Mouton de Gruyter. 311-355.
- [9] Grice, Martine & Stefan Baumann. 2002. Deutsche Intonation und GToBI. *Linguistische Berichte* 191, 267-298.
- [10] Grice, Martine, Stefan Baumann & Ralf Benz Müller. 2005. German Intonation in Autosegmental-Metrical Phonology. In: Jun, S. (ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press. 55-83.
- [11] Grosz, Barbara, Aravind Joshi & Scott Weinstein. 1995. Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics* 2 (21), 203-225.
- [12] Gussenhoven, Carlos. 1983. Focus, Mode, and the Nucleus. *Journal of Linguistics* 19, 377-417.
- [13] Gussenhoven, Carlos. 2002. Intonation and Interpretation: Phonetics and Phonology. *Proceedings 1st Int. Conference on Speech Prosody*. Aix-en-Provence. 47-57.
- [14] Kohler, Klaus. 1991. Terminal Intonation Patterns in Single-Accent Utterances of German: Phonetics, Phonology and Semantics. *AIPUK* 25, 115-185.
- [15] Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.
- [16] Pierrehumbert, Janet B. & Julia Hirschberg. 1990. The Meaning of Intonational Contours in the Interpretation of Discourse. In: Cohen, Philip R., Jerry Morgan & Martha E. Pollack, (eds.), *Intentions in Communication*. Cambridge: MIT Press. 271-311.
- [17] Prince, Ellen F. 1981. Toward a Taxonomy of Given-New Information. In: Peter Cole (ed.), *Radical Pragmatics*, New York: Academic Press. 223-256.
- [18] Selkirk, Elisabeth. 1984. *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- [19] Yule, George. 1981. New, Current and Displaced Entity Reference. *Lingua* 55, 41-52.