Accent Type and Givenness: An Experiment with Auditory and Visual Priming

Stefan Baumann[†], Kerstin Hadelich[‡]

† Institute of Phonetics, Saarland University

‡ Institute of Psycholinguistics, Saarland University

E-mail: baumann@coli.uni-sb.de, hadelich@coli.uni-sb.de

ABSTRACT

In this paper we address the question of whether the intonational marking of discourse referents in German varies according to their degree of activation. We distinguish three activation states corresponding to given, accessible, and new information (cf.[1]). In a perception experiment using synthesised sentences as auditory and drawings as visual primes, we investigated the effect of the mode of presentation on a referent's assumed activation state and the appropriateness of three different accent patterns for the marking of different activation states. Our results show differences in the priming effect on the activation state for the two modes of presentation: Auditorily primed referents seem to be activated to a higher degree than visually primed ones. Furthermore, it could be shown that the type of nuclear accent is a relevant cue for the perception of different degrees of givenness.

1. INTRODUCTION

In a recent cross-linguistic study on the intonational marking of textually given material, Cruttenden [2] found evidence against the assumption that given information universally is deaccented. In many languages, in particular Romance languages, deaccenting is at best one option among others (e.g. syntax) to establish discourse cohesion. However, for the West Germanic languages English and German a clear preference for deaccenting as a marker for given information could be confirmed.

Taking this general tendency in German for granted, the results of a psycholinguistic production experiment (picture description task, [3]) came as a surprise: Speakers only very rarely deaccented referents which had been visually presented as a prime. Rather, the primed referent tended to coincide with a nuclear accent in the linguistic realisation. Nonetheless, the prime led to a significant change in the syntactic structure which subjects used in their picture descriptions: When a primed referent occurred in a picture in the thematic role of patient, the number of patient first sentences (passives) increased compared to a baseline condition in which no prime was shown.[1] This suggests that a visually primed entity gets activated and can influence the choice of syntactic structure, but is not given enough to be deaccented.

These observations raise two questions. First, is there an accent on activated information in German, which can be distinguished from a "newness-accent", and is its use, in certain contexts, preferable to deaccentuation? Second, is there a difference in the intonational marking of an auditorily primed referent versus a visually primed one?

The theoretical framework for these questions will be discussed in section 2, followed by the description of a perception experiment in section 3, and the general discussion in section 4.

2. THEORETICAL BACKGROUND

2.1. The Notion of "Givenness"

Following Chafe [1], we regard givenness as a cognitive notion reflecting the activation status of a discourse referent which the speaker assumes to be present in the listener's consciousness at the time of utterance. Referents are understood as the ideas corresponding to entities of the external world in the minds of speakers and listeners. These entities consist of people, objects, and abstractions rather than events and states, the latter being less persistent in a person's consciousness and thus not serving as anchor points for new information over a larger stretch of discourse in the same way as referents do [1:79f.].

2.1.1. Degrees of Givenness

Givenness in terms of activation implies the assumption of different *degrees* of givenness. Chafe postulates three states of information in terms of the activation cost a speaker has to invest in order to transfer an idea from a previous state into an active state: *given, accessible*, and *new*. If a referent is already active in the listener's consciousness at the time of the utterance, it is *given*; if a referent gets activated from a previously semiactive state, it is *accessible*; if a referent gets activated from a previously inactive state, it is *new*. Lambrecht [4:109] further divides accessible information into textually,

---

inferentially and situationally accessible information, taking the nature of the context into account. He claims that referents which can be inferred from a textually given entity (inferentially accessible) and referents which are present in the surroundings of the discourse participants (situationally accessible) do not necessarily have to be in the peripheral consciousness of the listener, since he considers cognitive accessibility a *potential* for easy activation rather than a state of a referent in a person's mind. Thus, the question is not whether a referent is "objectively" active or inactive in a listener's mind, but whether a speaker assumes that a listener is willing and able to draw the intended inferences on the basis of the linguistic forms chosen by the speaker [4:105].

2.1.2. The Linguistic Form of Referring Expressions

Most studies on givenness concentrate on the appropriate forms on different linguistic levels (in particular phonology and morphosyntax) which a referring expression can take according to the assumed activation state of the referent. This is done either in the form of fine-grained taxonomies (e.g. [5]) or by using a gradual scale [6]. Other, more general approaches, combine parameters like shared knowledge and activation cost with the means of their linguistic manifestation (e.g. [7], [8]). Only few studies, however, are concerned with a closer look at the phonological form of referring expressions, which goes beyond a simple two-way distinction between accented and unaccented. One such English study is Brazil et al. [9], who claim that it is the *type* of the nuclear accent (falling vs. falling-rising) that distinguishes new from given information, contrary to the general assumption that given information is deaccented (at least in German and English).

2.1.3. Accent Type and Givenness

Taking up the idea of distinguishing different activation states by different types of accent, we are concerned with the distinction between new and accessible information, since Chafe claims that "both new and accessible information are expressed with accented full noun phrases, whereas given information is expressed in a more attenuated way." [1:75].

"Accessibility-accents", as opposed to "newness-accents", have been proposed both for English and German. For American English, Pierrehumbert and Hirschberg [10] claim in their compositional approach of intonational meaning that an H* pitch accent followed by a !H* (downstepped H*[2]) marks an inferrable item, and an H+!H* pitch accent highlights a predication which is already mutually believed by speaker and listener. We can subsume these accent types under accents marking accessible information. New information, on the other hand, is marked by an H*, whereas given information is assigned an L* accent (if accented at all). For German,

Kohler [12] proposed a so-called *early peak* contour, describing a pitch movement with a peak immediately preceding a considerably lower accented syllable, which is supposed to mark activated information. It is opposed to a medial peak on an accented syllable, which is said to mark newness. However, early peak refers to the marking of a higher-level semantic-pragmatic relation and is not meant to describe the activation state of a concrete referent at utterance-level. The early peak contour is taken to be equivalent to an H+L* or H+!H* pitch accent in GToBI, the medial peak to an (L+)H* pitch accent. Other approaches to assigning meanings to specific types of pitch accent in German largely concentrate on differences in sentence modality and Focus-Background structure (see [11] for a summary).

The perception experiment described in section 3 examines the relation between activation status and accent form in German.

2.2. The Establishment of Givenness in Spoken Discourse

Considering givenness in the context of spoken discourse, the question arises whether an entity's activation state (and thus its degree of givenness) is dependent on the mode in which it was introduced. There are at least two different ways in which the givenness of an entity may be established in (spoken) discourse: First, a referent may be given through the use of language, either explicitly (e.g. *a car* → *the car*) or implicitly (e.g. *my handbrake* → *the car*). Second, a referent may be non-linguistically given, e.g., through objects in the environment of the discourse participants.

It is often taken as a fact in psycholinguistic priming experiments that auditorily presented items lead to a greater priming effect on a subsequent target structure than visually presented ones. Nevertheless, these experiments were generally concerned with the syntactic structure of the output utterances. As mentioned above, however, a priming effect on the syntactic level (e.g. an increase of passive constructions) does not necessarily coincide with a priming effect on the phonological level. It is this effect on the *phonological* level that has to be examined in different priming conditions. In spoken discourse, theses conditions materialise in form of auditorily and visually available information. Chafe points out the difference between these two modes and postulates a so-called *echoic memory* [1:55], which suggests a higher degree of accessibility for auditorily given referents. This alleged difference in activation should be reflected in the (phonological) form of the referring expression a speaker chooses.

In the experiment described in the following section, we examine the role of mode of presentation and its influence on the referents' activation states by using both auditory (linguistic) and visual (non-linguistic) primes.

---

[2] The original notation has been translated into GToBI [11], the annotation scheme for German intonation used in the present study.

3. PERCEPTION EXPERIMENT

3.1. Hypotheses

The appropriateness of three different types of nuclear intonation patterns was tested in three different priming conditions. The primes varied in terms of activation degrees of the nuclear referents captured in the mode of presentation. We hypothesise that a referent is fully activated or given after having been presented as an auditory prime. Visually priming is expected to lead to a semi-active or accessible state of the primed referent. Finally, if no prime precedes an utterance, we regard all referents as previously inactive or new information.

As far as the intonational marking of the activation states is concerned, we hypothesise – in accordance with the studies mentioned in section 2.1.3. – that an H* pitch accent is the preferred marker of new information, and an H+L* pitch accent of accessible information, whereas deaccentuation is most acceptable with given information.

The hypotheses are summarised in Table 1.

given	accessible	new
auditory prime	visual prime	no prime
no accent	H+L*	H*

Table 1: Hypotheses of the relation between priming condition and nuclear accent type.

3.2. Setup

30 native speakers of German took part in the experiment. We presented 27 black-and-white drawings depicting actions between two entities (e.g. a parrot calling a waitress). These pictures were used in a previous psycholinguistic production experiment mentioned in the introduction. They had been pretested with respect to naming reliability, so that we could be sure that the descriptions would not be judged as inappropriate. With a one second delay, subjects heard a synthesised description of the depicted action (target sentence). The descriptions were produced by a diphone synthesis system [13] with GToBI input [11] using a male voice. The (morpho-) syntactic form of the target sentences was kept constant, i.e., we generated declarative sentences with a direct object (coded as a full definite noun phrase) in final position. Similarly, the accent type on the subject of the sentence (L+H*) was kept constant across conditions. The accent type of the direct object of the sentence was varied: It could either carry an H* pitch accent, an H+L* pitch accent or be deaccented (Ø, with the accent shifted onto the verb), as illustrated in Figure 1.

Each picture description was preceded by a context (prime) which could either be auditory, visual, or neutral. The auditory context consisted of a synthesised sentence using a female voice introducing the object of the following picture description, e.g. "Im Restaurant bedient eine Kellnerin" (lit.: *In the restaurant serves a waitress*).
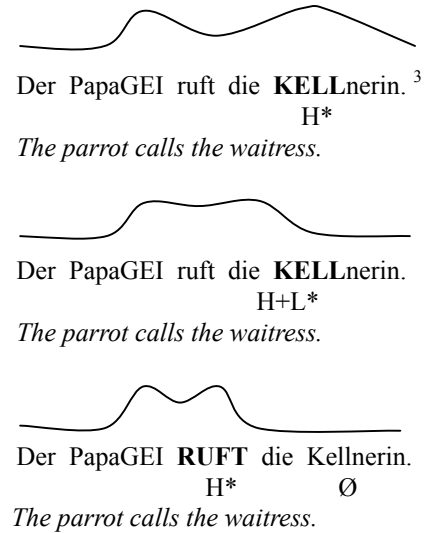

Der PapaGEI ruft die **KELL**nerin. [3]
　　　　　　　　　　　H*
The parrot calls the waitress.

Der PapaGEI ruft die **KELL**nerin.
　　　　　　　　　　　H+L*
The parrot calls the waitress.

Der PapaGEI **RUFT** die Kellnerin.
　　　　　　　H*　　　　Ø
The parrot calls the waitress.

Figure 1: Schematised intonation contours of the target sentences.

The prime sentences carried a nuclear accent on the object. Note that the relevant entity (in this case *the waitress*) occurred in the same surface position in both the prime and the target sentence and can thus be regarded as given. The visual context was a picture showing the relevant entity in a neutral posture. Subjects saw the visual prime for one second. In the neutral context condition neither a sound nor a picture was shown. In this no-prime condition the relevant entity in the target sentence was considered new.

As trained in four practice trials, participants were asked to judge the appropriateness of the intonation of the target sentence on a scale from 1 to 7. They were instructed to make their judgements regardless of the speech synthesis' segmental quality, the wording of the sentence, and the plausability of the displayed action.

3.3. Results

The data of three participants had to be excluded from the analysis due to technical problems. In order to compensate for inter-individual differences in the scorings, the scoring data of the remaining 27 participants were z-transformed such that each participant had a mean score of 0 and a standard deviation of 1. These z-scores entered the actual analyses.

Overall results showed a highly significant interaction between accent type (including deaccentuation) and priming condition ($F_{(4, 728)} = 27.82$; $p < .001$).

In Table 2, we summarise the significant overall results of the Anova and the results of the paired comparisons (Scheffé Posthoc) we conducted, sorted according to priming condition and type of nuclear pitch accent.

---

[3] Capital letters indicate accented syllables, bold face letters indicate nuclear accents. GToBI-notations are restricted to nuclear accents. The symbol Ø, which is not part of the GToBI annotation scheme, indicates lack of accent. It is only used for illustration purposes.

	Anova	Posthoc
auditory prime (given)	F(2, 239) = 26.68; p < .001	Ø > H+L* > H*
visual prime (accessible)	---	---
no prime (new)	F(2, 240) = 30.92; p < .001	H* > Ø H+L* > Ø

	Anova	Posthoc
Ø	F(2, 240) = 30,19; p < .001	auditory prime > visual prime > no prime
H+L*	---	---
H*	F(2, 231) = 19,49; p < .001	no prime > visual prime > auditory prime

Table 2: Significant relations between priming conditions and nuclear pitch accent type. The ">" symbol is to be interpreted as "significantly preferred over".

4. DISCUSSION

Our experiment clearly confirms the general assumption that new information is preferably marked by a pitch accent. However, there is no significant preference for the *type* of accent marking newness. There is, nevertheless, indirect evidence in favour of H*, since this accent type is significantly more acceptable in the no-prime (new) condition in comparison to the other conditions. There is no such effect with the other pitch accent type tested, H+L*.

The data further suggest that deaccentuation is most appropriate to mark given information, provided that an auditorily presented element *can* be called given. Indeed, there are good reasons to think so: First, there might be an activation support for spoken language due to echoic memory. Second, auditorily presented material is textually or linguistically given, which in other studies (e.g. [2]) has been taken to be activated information. Finally, the nature of the task might have a crucial influence: Since subjects were asked to evaluate the appropriateness of an *auditorily* presented target sentence, they may have been more sensitive to the auditory channel, which could have increased the referent's activation degree.

The auditory priming condition provides another significant difference: Pitch accent type H+L* was preferred over H* for marking the activated referents. This can serve as (at least indirect) evidence for the role of H+L* as an "accessibility-accent" or "activation-accent" rather than a "newness-accent", which is in line with our hypothesis.

The visual priming condition did not trigger a significant preference of pitch accent type in the target sentences, which indicates that the status of referents established by this mode of presentation is not as clear-cut as in the auditory mode. However, evidence that the two conditions are different is provided by the fact that H* is more acceptable and deaccentuation is less acceptable after visual than after auditory priming. We interpret this to mean: Visually presented referents are "less given" than auditorily presented referents.

REFERENCES

[1] W. Chafe, *Discourse, Consciousness, and Time*, Chicago/London, Univ. of Chicago Press, 1994.

[2] A. Cruttenden, "The Deaccenting of Given Information: a Cognitive Universal?", in G. Bernini (ed.), *The Pragmatic Organisation of Discourse*, Mouton, de Gruyter, in press.

[3] K. Hadelich, "Priming Planning, Production or Both." Poster presented on 8th AMLaP, Tenerife, Spain, 2002.

[4] K. Lambrecht, *Information Structure and Sentence Form*, Cambridge, CUP, 1994.

[5] M. Ariel, "The Function of Accessibility in a Theory of Grammar", *Journal of Pragmatics* 16, pp. 443-463, 1991.

[6] E. Hajicova, *Issues of Sentence Structure and Discourse Patterns*, Theoretical and Computational Linguistics, Vol. 2, Prague, Charles University, 1993.

[7] D.J. Allerton, "The Notion of 'Givenness' and its Relation to Presupposition and Theme", *Lingua* 44, pp. 133-168, 1978.

[8] E.F. Prince, "Toward a Taxonomy of Given-New Information", in *Radical Pragmatics*, P. Cole (ed.), New York, Academic Press, pp. 223-256, 1981.

[9] D. Brazil, M. Coulthard and C. Johns, *Discourse Intonation and Language Teaching,* London, Longman, 1980.

[10] J. Pierrehumbert and J. Hirschberg, "The Meaning of Intonational Contours in the Interpretation of Discourse", in *Intentions in Communication*, P.R. Cohen, J. Morgan & M.E. Pollack (eds.), Cambridge (Mass.), pp. 271-311, 1990.

[11] M. Grice, S. Baumann and R. Benzmüller, "German Intonation in Autosegmental-Metrical Phonology", in *Prosodic Typology: Through Intonational Phonology and Transcriptions*, S.-A. Jun (ed.), Oxford, OUP, in press.

[12] K.J. Kohler, "Form and Function of Intonation Peaks in German: A Research Project", AIPUK 25, 1991.

[13] M. Schröder and J. Trouvain, "The German Text-to-Speech Synthesis System MARY: A Tool for Research, Development and Teaching", Proceedings 4th Speech Synthesis Workshop, Pitlochry (Scotland), pp. 131-136, 2001.