

On the Perception of Intonationally Marked Givenness after Auditory and Visual Priming

Stefan Baumann^{*}, Kerstin Hadelich^{**}

Abstract

This paper deals with the intonational marking of different activation degrees of discourse referents. Following Chafe (1994), we distinguish three activation states corresponding to given, accessible, and new information. In a perception experiment using synthetic speech and pictures, we investigate whether the mode of presentation (auditory vs. visual) has an influence on a referent's assumed activation state and how appropriate different accent patterns are for different activation states. Our findings suggest that auditorily primed referents are activated to a higher degree than visually primed ones, and that the type of the nuclear accent is a relevant cue for the perception of different degrees of givenness.

1. Introduction

It is widely assumed for West Germanic languages like English and German that new information is marked by a pitch accent, while given information is deaccented (e.g. Ladd, 1996). This assumption has recently been confirmed by a cross-linguistic study on the intonational marking of textually given material (Cruttenden, in press). Cruttenden points out, however, that deaccenting given information is not a cognitive universal. In many languages, in particular Romance languages, deaccenting is at best one option among others (e.g. syntax) to establish discourse cohesion.

Expecting that given information is deaccented in German, the results of a psycholinguistic production experiment (picture description task; Hadelich, 2002) 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 a patient, speakers tended to use patient-first sentences (passives) in their descriptions rather than the canonical agent-first structure. This suggests that a visually primed entity gets activated and can influence the choice of syntactic structure, but is not given enough to get deaccented. These observations raise two questions:

a) 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?

b) What is the influence of the mode of presentation? In particular, 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. Givenness

2.1. Degrees of Givenness and their Linguistic Manifestation

When Halliday (1967) introduced the terms *given* and *new* in his model of information structure, he defined them as a dichotomy: Given information is presented by the speaker as being recoverable from the discourse context, new information is not. The relation of given vs. new is largely equivalent to what is often referred to as *Background* vs. *Focus*. Thus, it describes the internal partitioning of an informational unit, which is equivalent to an intonation unit, rather than the cognitive states of single discourse referents.

^{*} Institute of Phonetics, Saarland University, baumann@coli.uni-sb.de

^{**} Institute of Psycholinguistics, Saarland University, hadelich@coli.uni-sb.de

Somewhat deviating from Halliday and following Chafe (e.g. 1994), 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 "real-world" entities 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 (Chafe, 1994:79 f.).

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 (1994:109) 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 (Lambrecht, 1994:105).

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. Ariel, 1991, Givón, 1990, Gundel et al., 1993) or by using a gradual scale (Hajicova, 1993). Other, more general approaches, combine parameters like shared knowledge and activation cost with the means of their linguistic manifestation (e.g. Allerton, 1978, Prince, 1981). 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. (1980), 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).

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." (1994:75).

"Accessibility-accents", as opposed to "newness-accents", have been proposed both for English and German. For American English, Pierrehumbert and Hirschberg (1990) claim in their compositional approach of intonational meaning that an H* pitch accent followed by a !H* (downstepped H* ¹) 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 (1991) proposed a so-called *early peak* contour which is supposed to mark activated information, whereas a medial peak on an accented syllable marks newness. However, early peak refers to the marking of a higher-level semantic-pragmatic relation rather than to a concrete referent at the 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 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 Grice et al. for a summary).

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

¹ The original notation has been translated into GToBI (Grice et al., in press), the annotation scheme for German intonation used in the present study.

2.2. The Establishment of Givenness in Spoken Discourse

As we have seen in the last section by Lambrecht's distinction between textually, inferentially, and situationally accessible information, there are at least two different ways in which the givenness of an idea may be established – note that we are concerned with 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*). Implicit textual givenness is equivalent to Lambrecht's category of "inferentially accessible information". This type of reference has also been called *endophoric* (Halliday & Hasan, 1976:33). Second, a referent may be non-linguistically given, e.g., objects in the environment of the discourse participants. Halliday and Hasan called this type of reference *exophoric* (1976:33).

The question of interest for us is whether a linguistically given referent is evoked in a listener's consciousness to the same extent as a non-linguistically given referent. A crucial factor may be the mode of activation. While linguistic givenness materialises in spoken language, a non-linguistically given referent is, for instance, visually present. Chafe points out this difference when he claims that "sound remains briefly available to active consciousness even if it failed to enter that state while it was physically present." (1994:55). He calls this phenomenon *echoic memory*. It suggests a different quality of the processing of auditory and visual stimuli. Presumably, auditorily given referents are more easily accessible than visually given referents. This alleged difference in activation should be reflected in the form of the referring expression a speaker chooses. We assume – again - that it is the *phonological* form (i.e. the intonation pattern) of the referring expression that makes the difference.

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.

3. Perception Experiment

3.1. Hypotheses

This experiment was designed to test the appropriateness of three different types of nuclear intonation patterns 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 at the forefront of a listener's consciousness, i.e. fully activated or given, after having been presented as an auditory prime. We equate a visually primed referent with semi-active or accessible information. 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. – 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 and pretested with respect to naming reliability, so that we could be sure that the descriptions would not be judged as inappropriate. With one second delay, subjects heard a synthesised description of the depicted action (target sentence). The descriptions were produced using diphone synthesis with GToBI input (Schröder & Trouvain, 2001) and spoken with 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).

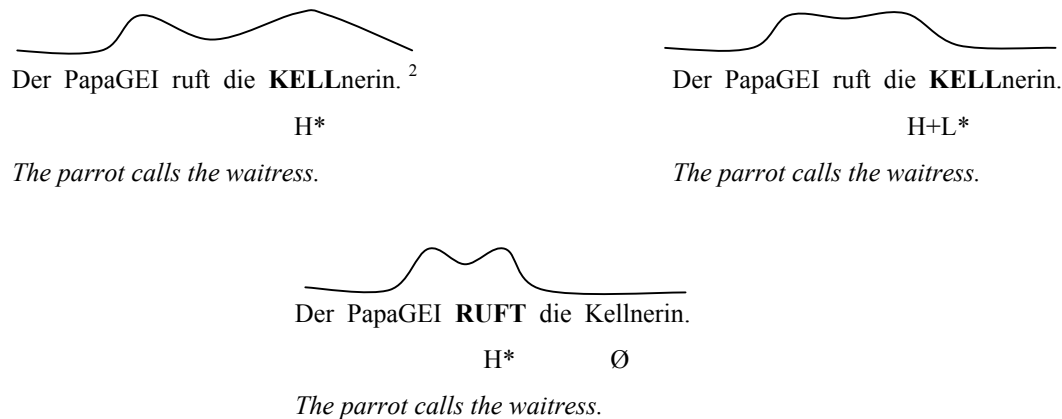


Figure 1: Schematised intonation contours of the target sentences

Each picture description was preceded by a context (prime) which could either be auditory, visual, or neutral (cf. Figure 2). The auditory context consisted of a synthesised sentence spoken by 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*). 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 plausibility 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 ($F(4, 728) = 27.82$; $p < .001$). However, the effect of accent pattern considerably varied within the different priming conditions. In the no-prime (newness) condition, both accent types were significantly preferred over deaccentuation ($F(2, 240) = 30.92$; $p < .001$), whereas the distinction between H* and H+L* was not significant. The visual-prime condition, which we assumed to trigger accessible referents, did not provide significant differences between the three accent patterns. In the auditory-prime (givenness) condition, however, the phonological differences in

² 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.

the target sentences was highly significant ($F(2, 239) = 26.68$; $p < .001$): Deaccentuation was the preferred marking of an already activated referent, as hypothesised. Also, H+L* turned out to be significantly more acceptable than H* for marking given information.

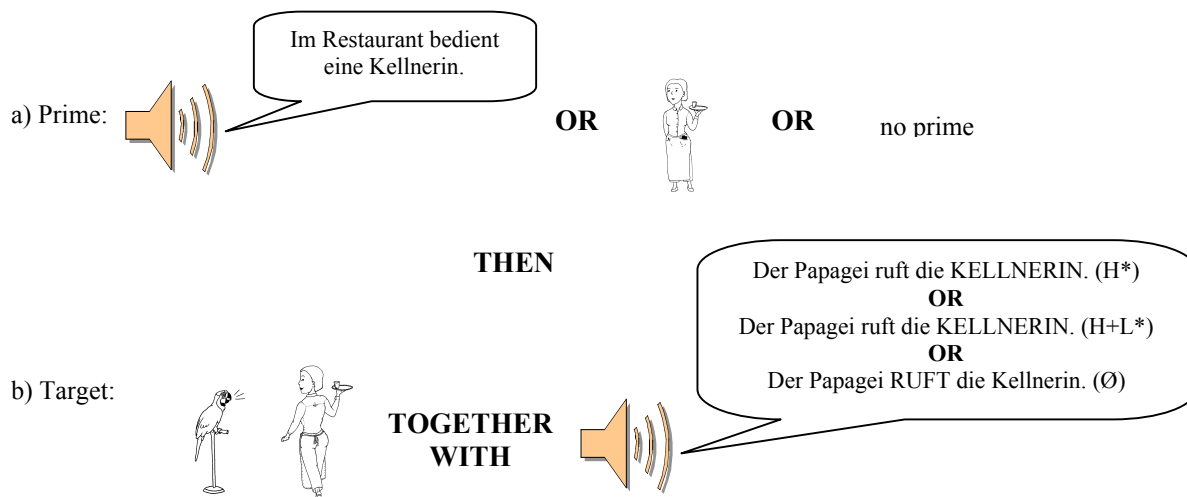


Figure 2: Procedure of the perception experiment

To clarify further the appropriateness of the individual accent patterns we also conducted an analyses across priming conditions. For the H* pitch accent as well as for deaccentuation, the influence of all three priming conditions had significant effects ($F(2, 231) = 19,49$; $p < .001$ and $F(2, 240) = 30,19$; $p < .001$; respectively). H* is significantly more acceptable after no prime than after a visual prime, and more acceptable after a visual prime than after an auditory one. Similarly, deaccentuation is preferred after auditory priming in comparison to visual priming, and after visual priming as opposed to no priming. On the other hand, the influence of the priming condition had no significant influence on the distribution of the H+L* pitch accent. These figures support our hypotheses that deaccentuation is most appropriate to mark given information and H* is most appropriate to mark new information.

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 suggest that auditorily presented referents can be considered given or activated information, since deaccentuation, which we take as a marker for givenness, is clearly preferred over accentuation. One reason might be the activation support of the echoic memory mentioned above. Another reason could lie in the dominant effect of *linguistically given* material, which was tested by Cruttenden who also found a strong tendency for deaccentuation in German. 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.

However, a simple equation of visually presented material with accessible information may not be valid. The status of a visually given referent is too vague, since no significant difference in its intonational marking could be found. Research in two directions would be desirable: First, we need more sophisticated experiments on visual givenness in a less artificial setup, including "real-world" objects instead of pictures. Second, we need experiments with textually accessible referents, in the sense of semi-active information due to previous mention that is either displaced or evoked as part of a given scenario or schema (Chafe, 1994:122). A better understanding of the nature of accessibility seems to be a necessary prerequisite for further studies on the intonational marking of information between the poles given and new.

Acknowledgement

This research is supported by the EC Project NECA IST-2000-28580. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

References

- Allerton D.J., 1978, "The Notion of 'Givenness' and its Relation to Presupposition and Theme", *Lingua* 44, p. 133-168.
- Ariel M., 1991, "The Function of Accessibility in a Theory of Grammar", *Journal of Pragmatics* 16, p. 443-463.
- Brazil D., Coulthard M. & Johns C., 1980, *Discourse Intonation and Language Teaching*, London, Longman.
- Chafe W., 1994, *Discourse, Consciousness, and Time*, Chicago/London, Univ. of Chicago Press.
- Cruttenden A., in press, "The Deaccenting of Given Information: a Cognitive Universal?", in G. Bernini (ed.), *The Pragmatic Organisation of Discourse*, Mouton, de Gruyter.
- Givón T., 1990, *Syntax: A Functional-Typological Introduction*. Vol. 2, Amsterdam, Benjamins.
- Grice M., Baumann S. & Benz Müller R., in press, "German Intonation in Autosegmental-Metrical Phonology", in *Prosodic Typology: Through Intonational Phonology and Transcriptions*, S.-A. Jun (ed.), Oxford, OUP.
- Gundel J., Hedberg N., & Zacharski R., 1993, "Cognitive Status and the Form of referring Expressions in Discourse", *Language* 69, p. 274-307.
- Hadelich K., 2002, "Priming Planning, Production or Both." Poster presented on 8th AMLaP, Tenerife, Spain.
- Hajicova E., 1993, *Issues of Sentence Structure and Discourse Patterns*, Theoretical and Computational Linguistics, Vol. 2, Prague, Charles University.
- Halliday M.A.K., 1967, "Notes on Transitivity and Theme in English", Part 2, *Journal of Linguistics* 3, p. 199-244.
- Halliday M.A.K. & Hasan R., 1976, *Cohesion in English*, London, Longman.
- Kohler K.J. , 1991, "Form and Function of Intonation Peaks in German: A Research Project", AIPUK 25.
- Ladd D.R., 1996, *Intonational Phonology*, Cambridge, CUP.
- Lambrecht K., 1994, *Information Structure and Sentence Form*, Cambridge, CUP.
- Pierrehumbert J. & Hirschberg J., 1990, "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.), p. 271-311.
- Prince E.F., 1981, "Toward a Taxonomy of Given-New Information", in *Radical Pragmatics*, P. Cole (ed.), New York, Academic Press, p. 223-256.
- Schröder M. & Trouvain J., 2001, "The German Text-to-Speech Synthesis System MARY: A Tool for Research, Development and Teaching", Proceedings 4th Speech Synthesis Workshop, Pitlochry (Scotland), p. 131-136.