



PARALLELISM, COHERENCE, AND CONTRASTIVE ACCENT

Mariët Theune

IPO, Center for Research on User-System Interaction,
P.O. Box 513,
5600 MB Eindhoven,
The Netherlands
e-mail: M.Theune@tue.nl

ABSTRACT

This paper discusses an experiment concerning the assignment of contrastive accents, i.e., accents used to indicate the presence of contrastive information. The experiment tested which of two existing approaches to the determination of contrastive information gives the most natural results. According to one approach the presence of 'alternative items' is the only condition for the assignment of contrastive accent; according to the other the presence of parallelism between sentences also is a condition. The experimental results indicate that the presence of alternative items combined with parallelism always triggers a preference for contrastive accent, whereas in the absence of parallelism, accent assignment seems to depend on the degree of coherence of the utterance.

1. INTRODUCTION

In spoken natural language, certain words are pronounced with more prominence than others; these words are said to be *accented*. In Dutch, as in other Germanic languages, accent functions (among other things) as a marker of information status: words expressing information that is *new* with respect to the discourse are usually accented,¹ whereas words expressing *given* information -- i.e., information that is available from the preceding discourse -- are usually unaccented [1, 3, 7]. Additionally, words expressing *contrastive* information are always accented, even if the information they express may be regarded as given [2, 8]. In this paper, the term *contrastive accent* is used to refer to accents that signal a contrastive relation between the accented item and a set of alternatives from which it should be distinguished.

When automatically generating spoken language, e.g., for use in information services over the telephone, it is important to generate the same accentuation patterns as

would be produced by human speakers, as this makes the output more natural-sounding and comprehensible. This means that a spoken language generation system should be able to determine the information status of the words and phrases in its output. So far, most research in this area has focused on the distinction between given and new information, whereas the question of which information is contrastive has received less attention. Only recently have some approaches to the determination of contrast been put forward, the most computationally oriented of which are those of Prevost [11] and Pulman [12].

Both Prevost and Pulman claim that the presence of contrastive accent on a word referring to a certain item (e.g., 'dog') is related to the previous mention of an *alternative item*, i.e., a different object of the same basic type (e.g., 'cat'). However, in the algorithm presented by Prevost, the mere presence of an alternative item is a sufficient condition for the assignment of contrastive accent, whereas Pulman distinguishes an additional condition, namely that the two sentences in which the items occur should be *semantically parallel*. Informally, this means that the two sentences should express a similar kind of event, e.g., 'The cat eats fish' and 'The dog eats meat'.

This paper discusses an experiment that was set up to test which of the following two hypotheses is correct. *Hypothesis I*: the presence of an alternative item triggers the assignment of contrastive accent. *Hypothesis II*: the presence of an alternative item plus semantic parallelism triggers the assignment of contrastive accent.

2. EXPERIMENT

The assumptions underlying the experiment are the following. Information is *given* if it has been expressed previously in the same discourse segment [3, 4, 8], and information is *new* if it is not given. Both given and new information can be *contrastive*. Words expressing information that is either new or contrastive are accented,² whereas words expressing given information are normally unac-

1. Accentuation not only depends on information status, but also on syntactic and lexical factors. For instance, function words such as determiners and prepositions generally are not accented, even if they might be regarded as expressing new information.

2. In this paper it is assumed that there is no phonological difference between contrastive accent and newness accent. See [10] for some discussion on this issue

cented, except if this information is contrastive. The effect of the different combinations of newness, givenness and contrast is represented in Table 1, where a plus sign indicates that the corresponding phrase is accented, and a minus sign indicates that it is not.

With respect to semantic parallelism, it is assumed that two sentences describe the same kind of event if they have the same or synonymous main verbs, and the arguments of these verbs are of the same type (these are the alternative items).

2.1 Method

Twenty subjects (14 male, 6 female, of different ages and backgrounds) were presented with twenty short texts in Dutch, displayed on a computer screen. They were instructed to first read each text, and then listen to two spoken versions of it, which could be played and replayed by clicking on two buttons. The subjects had to indicate for each text which of its two spoken versions they found the most natural sounding. They knew that the versions only differed with respect to the pronunciation of the last sentence, but they were not told the exact nature of this difference.

The spoken versions of the texts differed with respect to the accentuation of a target word in their final sentence: in one version (the ‘accented’ version), this word was accented, in the other version (the ‘unaccented’ version) it was not. The versions were generated using the speech synthesis system SPENGI [5].

2.2 Materials

The texts used in the experiment were constructed to test Hypotheses I and II, and consisted of two or three sentences. The first sentence of each text introduced a target item *X*. The last sentence of each text also contained a reference to *X*, which used the same wording. *X* was assumed to constitute given information by then. This was important for the experiment because, as Table 1 shows, words expressing given information are only accented if this information is contrastive. Thus, a subject’s preference for the accented version of a text could be interpreted as a preference for *contrastive* accent on the target word (the

	new	given
contrastive	+	+
not contrastive	+	-

Table 1: information status and accentuation.

final reference to *X*). If the target words expressed new information this interpretation would not follow, because new information is always accented.

The texts were divided into three categories, representing three different types of context for the target word. The texts in **Category I** contained a reference to an alternative item *Y*, preceding the final reference to *X*, and they showed no semantic parallelism between the sentence introducing *Y* and the sentence containing the final reference to *X*. This corresponds to the context specified in Hypothesis I. Category I contained ten texts, one of which consisted of two sentences.

The texts in **Category II** also contained a reference to an alternative item *Y*, and additionally they showed semantic parallelism between the sentence introducing *Y* and the sentence containing the final reference to *X*. This corresponds to the context specified in Hypothesis II. Category II contained seven texts, one of which consisted of two sentences.

Finally, the texts in **Category III** contained no reference to an alternative item. This category was not related to the hypotheses, but was added to test the assumption that the presence of an intervening sentence does not affect givenness of an item. It contained three texts, all of which consisted of three sentences.

The three categories are illustrated in Table 2, which shows a schematic representation of the texts in each category, together with an example text and its translation. In the schematic representations, the parallel parts of two sentences are represented by *A* and *A'* respectively, and the non-parallel parts are represented by ‘...’. In the example texts, the phrase ‘de tandarts’ (the dentist) refers to the target object *X*, and the phrase ‘de patient’ (the patient) refers to the alternative item *Y*.

I	1. <i>X</i> ... 2. <i>Y</i> ... 3. <i>X</i> ...	De tandarts stond naast de tandartsstoel. De patient had een erg slecht gebit. De tandarts was ontspannen.	The dentist was standing beside the dentist’s chair. The patient had very bad teeth. The dentist was relaxed.
II	1. <i>X</i> ... 2. <i>Y A</i> 3. <i>X A'</i>	De tandarts stond naast de tandartsstoel. De patient was erg nerveus. De tandarts was ontspannen.	The dentist was standing beside the dentist’s chair. The patient was very nervous. The dentist was relaxed.
III	1. <i>X</i> ... 2. ... 3. <i>X</i> ...	De tandarts stond naast de tandartsstoel. Het ging om een simpele klus. De tandarts was ontspannen.	The dentist was standing beside the dentist’s chair. It was going to be a simple job. The dentist was relaxed.

Table 2: schematic representation and example text plus translation for each category.

2.3 Results

The results of the experiment are shown in Table 3. For each category, the table indicates the mean number of preferences for the unaccented and for the accented versions of the texts in that category, and whether the difference between the two is significant using the binomial test ($p < 0.05$). As Table 3 shows, there was a significant preference for the accented text versions in Category II. For the other two categories, the difference between preferences for the accented or the unaccented versions was not significant.

2.4 Discussion

The results for Categories I and II indicate two things. First, the mere presence of an alternative item does not trigger a significant preference for contrastive accent. Second, the presence of an alternative item combined with semantic parallelism does trigger such a preference. In other words, Hypothesis I must be rejected, whereas Hypothesis II should be accepted. These findings might be explained as follows. Because the events that are described in the semantically parallel sentences of Category II are very similar, people feel the need to emphasize the differences between them by accenting the words that refer to the different participants in the events (i.e., the alternative items). On the other hand, the events described in the non-parallel sentences of Category I are not similar, and people feel no need to distinguish them by emphasizing the alternative items.

Although the above explanation accounts for the significant preference for accentuation in Category II, as well as for the absence of such a preference in Category I, it cannot explain the differences between the results for texts *within* Category I, as shown in Table 4 below. For about half of the texts in Category I there is a preference for the accented version, and for the other half there is a preference for the unaccented version. As Table 4 shows, there is a similar difference between texts within Category III, but not within Category II. Since the differences between the texts within Categories I and III cannot be explained by the presence or absence of an alternative item or parallelism, another factor must play a role here. In the next section it is argued that this factor may be *coherence*.

3. COHERENCE

One method to estimate the level of coherence of a text is *Centering Theory* [6]. In this theory, the notion of *backward looking center* (C_b) plays an important role. A C_b is a discourse entity that is mentioned in a sentence (S_n) of a text and which was also mentioned in the preceding sentence (S_{n-1}). It functions as a kind of link between the two sentences, thus establishing coherence between them. For example, in the sequence ‘The dog chased the cat. The cat chased a mouse,’ the C_b of the second sentence is the cat.

	unaccented	accented	significance ($p < 0.05$)
I	10.9	9.1	-
II	4.7	15.3	+
III	11.0	9.0	-

Table 3: mean number of preferences for unaccented and accented versions per category.

	I	II	III
accented version	4	6	1
unaccented version	4	0	2
neither version	2	1	0
total number of texts	10	7	3

Table 4: number of texts per category for which the majority of subjects preferred (i) the accented version, (ii) the unaccented version, and (iii) neither version.

If none of the entities mentioned in a sentence S_n have been mentioned in sentence S_{n-1} , then S_n has no C_b .³ In this case, the coherence between the two sentences is obviously minimal.

To determine the influence of coherence on accentuation, the texts of Categories I and III⁴ were divided into two basic coherence classes, based on the presence or absence of a C_b in their final sentence (note that this is only a *local* measure of coherence between the last two sentences of the text). Texts that did not have a C_b in their final sentence were classified as *incoherent* (IC), and texts that did have a C_b in their final sentence were classified as *coherent* (C). According to Centering Theory, some texts in the C class are more coherent than others, depending on the characteristics of the C_b and its predecessor (see [6]). However, these subtle differences in coherence level within the C class were ignored.

The accentuation preferences per coherence class are shown in Table 5. The table clearly shows that coherence of a text (i.e., the presence of a C_b in the final sentence) contributes to a preference for its unaccented version, and that incoherence (i.e., the absence of a C_b) contributes to a preference for the accented version. The

3. The presence of a C_b is assumed to be optional, as in [9].

4. The Category II texts were left out of consideration, because in this category (containing three C and four IC texts) there seems to be no relation between coherence and accentuation. As Table 4 shows, for all texts (except for one, where the versions scored equal) the accented version was preferred, regardless of coherence class. This suggests that the effect of semantic parallelism is strong enough to overrule any effects of coherence.

	accented version	unaccented version	neither version	total number of texts
IC	4	0	1	5
C	1	6	1	8
total number of texts	5	6	2	13

Table 5: number of texts per coherence class (combined for Categories I and III) for which the majority of subjects preferred (i) the accented version, (ii) the unaccented version, and (iii) neither version.

relationship between coherence class and accentuation preference is significant ($\chi^2 = 8.05$, $\alpha = 0.02$).

The perceived relationship between coherence and accentuation can be explained if we assume that a sentence which lacks a C_b , and is therefore not explicitly linked to the preceding sentence, is perceived by the subjects as the initial sentence of a new *discourse segment*, which shifts the listener's attention to another topic.⁵ Since givenness is generally assumed not to range across discourse segments, the information expressed in such a segment-initial sentence will once again be regarded as new, and the corresponding phrase will be accented. This means that the general preference for the accented versions of the IC texts in Categories I and III may be attributed to newness of the target object, not contrast.

4. CONCLUSION

The results of the experiment discussed in this paper indicate that Hypothesis I, as stated in the introduction, must be rejected and that Hypothesis II should be accepted: semantic parallelism triggers contrastive accent, but the mere presence of an alternative item does not. For texts that showed semantic parallelism between the two last sentences, there was a significant preference for the accented versions, whereas for texts without semantic parallelism there was no clear preference for either the accented or the unaccented versions. For the latter type of texts, the preference for one or the other version seemed to depend on their level of coherence rather than on the presence or absence of an alternative item. It would be interesting to further explore the connection between coherence and accentuation in a more elaborate experiment, set up specifically for this purpose.

5. ACKNOWLEDGEMENTS

I wish to thank Emiel Kraemer, Jan Landsbergen, Marc Swerts, Jacques Terken and Raymond Veldhuis for helpful comments on this paper and/or the experiment; Danny

5. Based on a corpus study, Passonneau [11] claims that discourse segmentation is not correlated with centering transitions, but with overspecification of noun phrases. However, the results of the current experiment do not show such a correlation.

Kersten for technical support in setting up the experiment; and all the subjects who participated in it. This research was carried out within the Priority Programme Language and Speech Technology (TST), which is sponsored by NWO (The Netherlands Organization for Scientific Research).

6. REFERENCES

- [1] Brown, G. Prosodic structure and the given/new distinction. In A. Cutler and D.R. Ladd (Eds.), *Prosody: Models and Measurements*, 67-77. Berlin: Springer Verlag, 1983.
- [2] Chafe, W.L. Language and consciousness. *Language*, 50: 111-133, 1974.
- [3] Chafe, W.L. Givenness, contrastiveness, definiteness, subjects, topics and points of view. In C.N. Li (Ed.), *Subject and Topic*, 25-55. New York: Academic Press, 1976.
- [4] Deemter, K. van. What's new? A semantic perspective on sentence accent. *Journal of Semantics*, 11:1-31, 1994.
- [5] Gigi, E. and L.Vogten. A mixed-excitation vocoder based on exact analysis of harmonic components. *IPO Annual Progress Report*, 32:105-110, 1997.
- [6] Grosz, B., A. Joshi, and S. Weinstein. Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2): 203-225, 1995.
- [7] Halliday, M.A.K. Notes on transitivity and theme in English. *Journal of Linguistics*, 3:199-244, 1967.
- [8] Hirschberg, J. Using discourse context to guide pitch accent decisions in synthetic speech. In G. Bailly, C. Benoît and T.R. Sawallis (Eds.), *Talking Machines: Theories, Models and Designs*, 367-376. Amsterdam: Elsevier Science Publishers, 1992.
- [9] Kameyama, M. A property-sharing constraint in centering. *Proceedings of the 24th Annual Meeting of the Association of Computational Linguistics, ACL-86*, 200-206, 1986.
- [10] Kraemer, E. and M. Swerts. Reconciling two competing views on contrastiveness. *Proceedings of the 5th International Conference on Spoken Language Processing, IC-SLP-98*, 1299-1302, 1998.
- [11] Passonneau, R. Interaction of discourse structure with explicitness of discourse anaphoric noun phrases. In M. Walker, A. Joshi and E. Prince (Eds.), *Centering Theory in Discourse*. New York: Oxford University Press, 1998.
- [12] Prevost, S. An information structural approach to spoken language generation. *Proceedings of the 34th Annual Meeting of the Association of Computational Linguistics, ACL-96*, 294-301, 1996.
- [13] Pulman, S. Higher Order Unification and the interpretation of focus. *Linguistics and Philosophy*, 20:73-115, 1997.