



DEACCENTUATION AND PERSISTENCE OF GRAMMATICAL FUNCTION AND SURFACE POSITION

Julia Hirschberg

Jacques Terken

*AT&T Bell Laboratories
Murray Hill NJ 07974 USA*

*Institute for Perception Research
5600 MB Eindhoven, The Netherlands*

Abstract

DEACCENTUATION is commonly explained as a consequence of the GIVENNESS of the information referred to. However, speakers often accent GIVEN information, so that a more complex explanation of deaccentuation is required. This work investigates the contribution of persistence of grammatical role and surface position to the occurrence of deaccentuation. We report a production experiment in which these features were varied systematically. The results indicated that persistence of grammatical role and surface position both contributed to deaccentuation.

Keywords: repairs, disfluencies

1. Introduction

A speaker's DEACCENTUATION of a particular word or phrase in a discourse has often been explained as a consequence of the information status of the item referred to: items that are ACCENTED, or made prosodically prominent, tend to represent GIVEN, or old, information in the discourse [2, 1]. Items that are accented tend to represent NEW information. However, speakers sometimes accent GIVEN information. Why speakers choose to accent some GIVEN items and deaccent others remains an open question. In this paper, we investigate some possible explanations for this behavior, in the persistence (or lack thereof) of grammatical function and surface position of GIVEN items in a discourse.

2. The Experiment

To investigate the possible contribution of persistence of grammatical function and surface position to a speaker's tendency to accent or deaccent GIVEN information, we compared speakers' accentuation behavior with respect to a referring expression x under the following scenarios: a) the referent of x has **not** been mentioned in the immediate context; b) the referent of x has been mentioned in the immediate context, but with a **different** grammatical function and/or surface position; and c) the referent of x has been mentioned in the immediate context, with the **same** grammatical function and/or surface position. For comparison with prior studies, we operationally define GIVEN as "mentioned in the prior discourse" and NEW as "not mentioned in the prior discourse". By comparing these conditions, we hoped to be able to determine the contributions of these factors and of GIVENNESS itself to speakers' tendency to deaccent referring expressions. In addition, by manipulating persistence of grammatical function and surface position independently, we hope to assess the relative effect of these factors.

We chose an elicitation technique for our study, presenting speakers with visual materials and asking them to describe

changes in the observed scene. We manipulated properties of the visual materials in such a way that we obtain the same descriptions in different contexts, thus allowing for comparison both within and between speakers. Despite the degree of predictability this technique produces, the task fulfills the requirement that the speaker has to construct each utterance. This technique has been applied before in the study of prosody by [6] and by [5].

Ten male speakers of American English participated as volunteers. Each was seated in a recording booth in front of a workstation screen displaying a series of changes in simple configurations of geometrical shapes. The task was to describe these changes. By manipulating the objects involved in the events, the role each object played in the event, and the way in which events were sequenced, we obtained utterances in which grammatical function, surface position, and information status (GIVEN-NEW) were systematically varied as desired.

Each subject described 75 scenarios, each consisting of a sequence of four events. The screen was cleared between scenarios, so that each scenario constituted a separate discourse. Only the subject's description of the last event in each scenario was analyzed. The previous events and their descriptions served only as context for this fourth event, which we will term the TARGET UTTERANCE; the GIVEN item(s) in that utterance we term the TARGET EXPRESSION(S).

All visual displays contained 9 objects: a ball, a cone, a cross, a star, a line, a box, a diamond, a rectangle, and a triangle. For ease of reference, and to discourage pronominal reference, each object was orthographically labeled. Objects could participate (as agent or patient) in three possible events: (1) an object could move towards another object, (2) an object could cover another object, and (3) an object could push another object against a third object. Typical speaker descriptions of these events were "The ball touches the cone", "The ball covers the cone", and "The ball pushes the cone against the cross." All three utterances contain a grammatical subject and a direct object; the third one also contains the object of a preposition. Each referring expression can be classified as either GIVEN or NEW in terms of the preceding context. The events in each scenario were designed so that GIVEN items would be described with the same grammatical function and surface position in all three of the context utterances. Below is a typical speaker description:

The ball touches the cone.
The ball touches the cross.
The ball touches the diamond.
The star touches the ball.

Here the target (GIVEN) expression in the target utterance is *the*

ball, which is the agent of the 'touching' event described in the first three scenes of the scenario and is described in subject position; by the second utterance it is also GIVEN, in the sense of "previously mentioned in the context". But in the fourth scene, *the ball*, while still GIVEN, is the patient of the 'touching' event, and is described in object position.

By varying events and the roles of objects in them across scenarios, the grammatical function of the target expression was orthogonally combined with its grammatical function in the context, yielding a 3x3 matrix with grammatical function in target utterance (SUBJ, DIRECT OBJ and PP-OBJ) and grammatical function in the context (SUBJ, DIRECT OBJ and PP-OBJ). The scenarios were constructed so that each target utterance was combined with three different contexts, so that the possible contribution of properties of the current utterance to accentuation was kept constant across conditions. (We will identify conditions in the form 'A→B', where 'A' gives the function in the prior context and 'B' gives the function in the target utterance.) In the SUBJ→PP-OBJ or DIRECT OBJ→PP-OBJ conditions, where the grammatical function of the target in the context was SUBJ or DIRECT OBJ and its grammatical function in the target expression was PP-OBJ, the preceding utterances in the scenario contained no PP-OBJ. Also, in the conditions where the target's grammatical function was SUBJ or DIRECT OBJ, the target utterances contained no PP-OBJ. As a result, in the conditions DIRECT OBJ→PP-OBJ and PP-OBJ→DIRECT OBJ, the target expression was in sentence-final position, both in the target utterance and in the context. This permitted an evaluation of the contributions of persistence of surface position and grammatical function. In each cell of the 3x3 design we had five scenarios, for a total of 45 target expressions per subject. In addition, the same target descriptions were combined with different contexts, in order to create two control conditions: one an ALL NEW condition, where the target expression was not mentioned in the prior context; one an ALL GIVEN condition in which the subject and direct object of the third event were also subject and direct object in the fourth (e.g. "The star touches the ball. The star covers the ball."), so that only the verb represented NEW information. There were fifteen scenarios for each control condition.

Each subject was treated individually and received oral and written instructions. He was instructed that the experiment was aimed at improving the way computers provide spoken information to humans. He was informed of the basic structure of the experiment, the possible events in the display, and the format of the scenarios, and was given examples of possible descriptions. He was introduced to a listener, who he was told would verify his descriptions relative to the events in another display but would provide no feedback; he was not, however, allowed to see the listener during the experiment. In the oral explanation, the subject was told to imagine he was telling a story or describing a movie he had seen. During the experiment, both subject and experimenter were in a relatively soundproof booth and seated before a SUN 361 workstation. Following the subject's description of an event in the display, the experimenter triggered the display of the next event. All events involved a change of the relevant object's or objects' position in a number of steps to a new position. The experiment was preceded by ten practice scenarios. The actual materials were presented in

Table 1: Average Accentedness of GIVEN Target Expressions as a Function of Grammatical Role in Target and Context

		Function in target utterance		
		SUBJ	DIRECT-OBJ	PP-OBJ
Function	SUBJ	2.1	3.6	3.2
in	DIRECT-OBJ	3.3	0.6	1.6
context	PP-OBJ	3.0	1.4	0.7

Table 2: Average Accentedness of SUBJ and DIRECT-OBJ in Control Conditions

		Function in current utterance	
		SUBJ	DIRECT-OBJ
ALL GIVEN		0.6	0.2
ALL NEW		3.3	3.6

Table 3: Average Accentedness of NEW Non-Target Expressions in Target Utterances as a Function of Grammatical Role

		Function in target utterance	
		SUBJ	DIRECT-OBJ
		3.7	3.8

a different random order for each subject, to balance for order effects. The speech was recorded by means of a Bruel and Kjaer microphone 2231, stored on DAT (sampling frequency 48 kHz), and downsampled to 12 KHz.

3. Results and Discussion

The target utterances for each subject were segmented from their contexts and subjected to two kinds of analysis: perceptual analysis, to determine the location of pitch accents, and pitch analysis, for measurement of acoustic features of perceived accents. The perceptual analysis was performed by the two authors independently, in such a way that they had no information about the context of a target description or the condition to which a target item belonged. For each NP in the target utterance, each judge indicated whether the object named was accented, unaccented, or of indeterminate accent status. These judgments were mapped onto a two point scale, from '0' for unaccented to '2' for accented. Summing the scores for the two judges gave each an *accentedness score* between 0 and 4. There was good agreement between our judges, with instances of full agreement at 81.3% and full disagreement (accented vs. deaccented) at only 7.3%. For the most part, our main results are based on the clear absence or presence of accents, ruling out the possibility that speakers deaccentuate or accentuate less clearly in some conditions than in others.

Pitch measurements were then obtained for all target utterances using a pitch tracker written by David Talkin. The resulting pitch tracks were visually inspected to determine pitch maxima for all referring expressions marked by a pitch accent; pitch maxima were generally measured at the place of maximum amplitude within the stressable syllable, except when the type of pitch accent realized on a referring expression was either H+L* or L*+H (in Pierrehumbert's system [4]), when the pitch maximum was measured at the high target of the accent

The main results of our experiment are presented in Tables 1-2. The results for individual conditions are expressed as num-

bers on our accentedness scale. For each condition, the means of these numbers were taken across replications and speakers. Table 1 contains the accentedness scores for GIVEN target expressions in target utterances. Table 3 contains the results for NEW, non-target expressions in the target utterances (This table contains data for SUBJ and DIRECT-OBJ only, since PP-OBJs in target descriptions were always target expressions.) Table 2 contains data for control conditions. Target utterances in these conditions contained only SUBJ and DIRECT-OBJ expressions. In some control conditions, both expressions were GIVEN and shared grammatical function and surface position with the previous utterance (ALL GIVEN). In others, both expressions were NEW.

Table 1 shows clear differences in degree of accentedness for the different context conditions. To evaluate these, the data were analyzed with an analysis of variance procedure; for each cell in the 3x3 matrix, degrees of accentedness were computed by subject. These data were used as data points for the analysis of variance, with grammatical function in the target utterance and grammatical function in the prior context as fixed factors and subjects as the replication factor. The two main effects were both significant: for function in prior context, $F(2,18) = 13.87, p = .0002$; for function in target utterance $F(2,18) = 21.97, p < .00005$. Also, the interaction was significant: $F(4,36) = 28.5, p < .00005$. That is, while both grammatical function in prior context and grammatical function in target utterance had significant effects on the accentedness of the target expression, the effect of grammatical function in context was not constant over different grammatical functions in the target utterance. The significant interaction makes the independent discussion of the effect of 'grammatical function in prior context' and 'grammatical function in target utterance' less meaningful, because the effect of the former effect is not the same for all grammatical functions in the target utterance. So, we looked at the simple main effects.

For conditions in which the target expression has been realized as grammatical subject in the prior context, the case of persistence of grammatical function to the target expression (SUBJ→SUBJ) contrasts with cases of change of grammatical function (SUBJ→DIRECT-OBJ and SUBJ→PP-OBJ); the SUBJ→SUBJ case shows a relatively low value (2.1) on the accentedness scale the other two show high values (3.6 and 3.2, respectively). A somewhat different effect is found for objects. For conditions in which the target expression was introduced in the context as DIRECT-OBJ, two conditions show relatively low values on the accentedness scale (DIRECT-OBJ→DIRECT-OBJ, .6, and DIRECT-OBJ→PP-OBJ, 1.6); in contrast, the DIRECT-OBJ→SUBJ condition shows a high accentedness value (3.3). Similarly, for conditions in which the target was introduced in the context as PP-OBJ, the conditions PP-OBJ→PP-OBJ and PP-OBJ→DIRECT-OBJ exhibit low values on the accentedness scale (.7 and 1.4, respectively), while the condition SUBJ→PP-OBJ exhibits a high value (3.0). That is, in cases where the target expression was initially either DIRECT-OBJ or PP-OBJ, persistence of grammatical function is associated with very low accent values for the target expressions. It may be noted, however, that change of function from DIRECT-OBJ to PP-OBJ, or vice versa, is associated with higher, but still fairly low accent values —

lower even than the value for the SUBJ→SUBJ condition, while a change in grammatical function from either DIRECT-OBJ or PP-OBJ in context to SUBJ in target utterance is associated with high accent values. One might in fact conclude from Table 1 that speakers deaccent GIVEN items when there is persistence of grammatical function from prior context to target — although this tendency is weaker for subjects than for objects.

Furthermore, there appears to be a contribution of persistence of surface position of the target expression from prior context to target utterance. A comparison of conditions in which there is persistence of surface position but not of grammatical function with conditions in which there is persistence of grammatical function **and** surface position supports this hypothesis. Conditions exhibiting persistence of surface position but not grammatical function (DIRECT-OBJ→PP-OBJ and PP-OBJ→DIRECT-OBJ) have higher values on the accentedness scale (1.6 and 1.4, respectively) than those exhibiting persistence of grammatical function — DIRECT-OBJ→DIRECT-OBJ (0.6) and PP-OBJ→PP-OBJ (.7); post-hoc analyses show that these differences are significant for DIRECT OBJ $F(1,9) = 9.7, p = .01$, but not for PP-OBJ, $F(1,9) = 2.8, p = .13$. These scores are still considerably lower than conditions which have a change both of grammatical function and surface position (DIRECT-OBJ→SUBJ, 3.3, PP-OBJ→SUBJ, 3.0, SUBJ→DIRECT-OBJ, 3.6, and SUBJ→PP-OBJ, 3.2). Thus, it appears that persistence of surface position contributes to deaccentuation — independent of the contribution of persistence of grammatical function. Since the two factors were not manipulated independently, we cannot investigate whether the two effects are independent and additive.

Note that the simple hypothesis that GIVENNESS (defined as "prior mention") explains deaccentuation can be ruled out for the data presented in Table 1 by comparing some conditions with the ALL NEW condition in Table 2. If we compare conditions for which there is no persistence of grammatical function or surface position in Table 1 (DIRECT-OBJ→SUBJ and SUBJ→DIRECT-OBJ) with the SUBJ and DIRECT-OBJ conditions for ALL NEW utterances in Table 2, we find no significant difference in accentedness ($F(1,9) = .71$). This confirms our original observation that the notion of GIVENNESS, without additional qualifications, is insufficient to explain subjects' deaccentuation behavior.

The interaction between 'grammatical function in context' and 'grammatical function in target utterance' explains the significant main effects only in part. Another, secondary effect is more surprising. If we look at the diagonal in Table 1 of conditions exhibiting persistence of grammatical function (SUBJ→SUBJ DIRECT-OBJ→DIRECT-OBJ, and PP-OBJ→PP-OBJ), we note that subjects are more frequently accented (2.1 on the accentedness scale) than direct objects (.6) or prepositional objects (.7). This is surprising, since it has often been argued that subjects are more likely to be deaccented than other expressions [3].

We noted above that there was full agreement between judges about whether or not an expression was accented in over 80% of cases; also, frequency distributions for the accentedness scale were uni-modal or bi-modal. Thus, an accentedness score near 2.0 for a GIVEN condition corresponds to approximately half of the relevant cases being accented and half of them being unaccented. Such a bi-modal distribution can be found in several

Table 4: Average Accentedness for Target Expressions For Individual Speakers in Three Conditions

Speaker	(a)	(b)	(c)
	same position same function	same position diff function	diff position diff function
1	0.3	0.4	3.0
2	1.1	1.2	3.4
3	0.0	3.2	3.6
4	1.8	2.6	2.8
5	1.1	1.8	2.7
6	1.0	1.5	3.9
7	0.0	0.4	3.6
8	0.5	2.0	3.7
9	0.0	1.0	2.8
10	0.5	1.0	4.0

cells of Table 1, including SUBJ→SUBJ, PP-OBJ→DIRECT-OBJ and DIRECT-OBJ→PP-OBJ. In order to determine whether this bi-modality can be explained as resulting from systematic differences between speakers, we inspected accentedness scores for individual speakers.

For the SUBJ→SUBJ condition, inspection of the data for individual speakers revealed no clear evidence of subgroups of speakers employing different strategies. Instead, there was a gradual transition from speakers who predominantly deaccented subjects to speakers who predominantly accented subjects in SUBJ→SUBJ condition. For PP-OBJ→DIRECT-OBJ and DIRECT-OBJ→PP-OBJ, however, there was evidence of differences in strategy by speaker group. In order to trace individual strategies, three different numbers were computed for each speaker: an average score for conditions in which the target expression had the same grammatical function and surface position in the target utterance and in the prior context (DIRECT-OBJ→DIRECT-OBJ and PP-OBJ→PP-OBJ), an average score for conditions in which surface position was constant but grammatical function was not (PP-OBJ→DIRECT-OBJ and DIRECT-OBJ→PP-OBJ), and an average score for conditions in which both grammatical function and surface position differed (SUBJ→DIRECT-OBJ and SUBJ→PP-OBJ). Again, these scores are computed on a five-point scale of accentedness, from 0 (deaccented) to 4 (accented). These scores are shown in Table 4. The data shown in Table 4 suggest that speakers fall into three subgroups with respect to their accenting behavior. For speakers 3 and 4, there is a large difference between {same surface position–same grammatical function} conditions on the one hand, and {same position–diff function} and {diff position–diff function} on the other. Apparently, for these speakers, “same function” is a necessary condition for deaccentuation. For speakers 1 and 2, there is a large difference between {same position–same function} and {same position–diff function} on the one hand, and {diff position–diff function} on the other. Apparently, for these speakers “same position” is a sufficient condition for deaccentuation. The remaining speakers appear to differentiate between all three situations. Apparently, for these speakers “same position” and “same function” are more or less independent conditions; the more the conditions

influencing deaccentuation are fulfilled, the more deaccentuation occurs.

4. Conclusion

Our findings confirm previous results showing that the GIVENNESS of an expression, where GIVENNESS is defined simply as “previous mention in context”, is not a sufficient condition for deaccentuation of that expression. Instead, additional requirements must be fulfilled, having to do with the surface position and grammatical function of an expression in the current utterance compared with its position and function in the prior context. The likelihood that a speaker will accent an expression representing GIVEN information thus depends on the relation between the expression in the current utterance and the antecedent expression in the context. Our main finding is that an expression will be deaccented if it has the same grammatical role as the antecedent expression in the immediate context (“persistence of grammatical function”). In addition, if the expression has a different grammatical function than the antecedent expression but occupies the same surface position (“persistence of surface position”), it will also be deaccented, but less frequently. And, if there is a change in both grammatical function and surface position from one utterance to the next, an expression is as likely to be accented as it would be if it represented NEW information.

Despite the importance of persistence of grammatical function and surface position for accent decisions, however, we also found that these effects are not constant across speakers. For some, persistence of grammatical role seems to be at least a necessary condition for deaccentuation; for others, persistence of surface position seems to be a sufficient condition; for still other speakers (the majority in our sample) persistence of position and of grammatical role seem to make independent and additive contributions.

References

- [1] W. Chafe. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li, editor, *Subject and Topic*, pages 25–55. Academic Press, New York, 1976.
- [2] M. A. K. Halliday. Notes on transitivity and theme in English, part 2. *Journal of Linguistics*, 3:199–244, 1967.
- [3] S. G. Nooteboom and J. Terken. What makes speakers omit pitch accents?: An experiment. *Phonetica*, 39:317–336, 1982.
- [4] J. B. Pierrehumbert. *The Phonology and Phonetics of English Intonation*. PhD thesis, Massachusetts Institute of Technology, September 1980. Distributed by the Indiana University Linguistics Club.
- [5] M. Swerts and R. Collier. On the controlled elicitation of spontaneous speech. *Speech Communication*, pages 463–468, 1992.
- [6] J. M. B. Terken. *Use and Function of Accentuation: Some Experiments*. PhD thesis, University of Leiden, Helmond, Neth, 1985.