DEACCENTUATION OF WORDS REPRESENTING ‘GIVEN’ INFORMATION: EFFECTS OF PERSISTENCE OF GRAMMATICAL FUNCTION AND SURFACE POSITION

JACQUES TERKEN
Institute for Perception Research, Eindhoven

and

JULIA HIRSCHBERG
AT&T Bell Laboratories, Murray Hill, New Jersey

The absence of intonational prominence on a referring expression (deaccentuation) is commonly explained as a consequence of the GIVENness of the discourse entity referred to – the fact that it represents old information in the discourse. However, speakers sometimes use accented expressions to refer to such GIVEN entities, so that GIVENness is not a sufficient explanation for deaccentuation. It has also been suggested that speakers tend to express GIVEN entities as grammatical subjects and to mention them early in the utterance. The present work investigates the contributions of grammatical role and surface position to the occurrence of deaccentuation in English. An experiment is reported in which speakers produced descriptions of visual materials, where the content of the materials was manipulated so that successive descriptions contained coreferential expressions, and grammatical role and surface position varied systematically. The results indicate that persistence of grammatical role and surface position from one utterance to the next both contribute to deaccentuation. Some implications for the way in which listeners may link referring expressions to entities which are already available from the context are discussed.

Key words: prosody, accentuation, grammatical function, surface position

INTRODUCTION

The distribution of accents in a discourse has often been explained with respect to the information status of the entities being mentioned: Words or word groups expressing GIVEN information tend to be deaccented, whereas words (or word groups) expressing

* This research was conducted while the first author was visiting AT&T Bell Laboratories, Murray Hill, NJ, during the spring of 1992. The visit was made possible by a grant from the Netherlands Organization for Scientific Research (NWO), Grant R 30-351, and additional funding from AT&T Bell Labs. and the Institute for Perception Research. The authors wish to thank Jan van Santen, Ananth Sankar and René Collier for useful discussions, and three anonymous reviewers for comments on a draft version.

Address correspondence to Jacques Terken, Institute for Perception Research, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, e-mail: terken@prl.philips.nl
NEW information tend to be accented (Chafe, 1976; Clark and Clark, 1977; Prince, 1981, 1992). The term "accent" is used here to mean "intonational prominence". In English, it is characterized by a conspicuous pitch change in or near the lexically stressed syllable of the word (Pierrehumbert, 1980; 't Hart, Collier, and Cohen, 1990)\(^1\) In addition, duration and amplitude of the stressed syllable are usually increased by accentuation (Fry, 1955, 1958; Lehiste, 1970). Furthermore, the spectral characteristics of the vowel in the stressed syllables of accented and unaccented words are often different (Van Son and Pols, 1992).

The association of accent with information status can be illustrated by examples such as shown in (1).

(1a) There are MOVIES and there are GOOD movies.

(1b) I bought a POODLE because I've always WANTED a dog.

In (1a), the word "movies" is deaccented on second mention, presumably because the concept has already been introduced into the discourse. In (1b), the word "dog" can be deaccented because the concept "dog" has been evoked by reference to a member of the class of dogs, the purchased poodle; the information status of "dog" is also GIVEN. Thus, GIVEN has been defined as "mentioned in the previous discourse or inerrable from the information mentioned previously".

Despite the existence of such persuasive examples in support of the GIVENness explanation of deaccentuation, there are also abundant counter-examples. Some are presented in (2).

(2a) There are MOVIES and there are MOVIES.

(2b) JOHN HIT BILL and then BILL CRUSHED JOHN.

(2c) President Clinton has arrived for a two-day visit in Moscow.

The PRESIDENT, who first visited Prague, ...

(2d) A: Why did you miss the party?

B: My MOTHER got sick.

\(^1\) Excluding cases of accent placement on all lexically unstressed syllables due to metalinguistic contrast (e.g., "I didn't say DEgrade, I said UPgrade.") or stress shift in complex nominals (e.g., "FIFteen CHInese JUGglers"). Here and below, capitals mark the accented syllables.
In (2a) it seems unlikely that any native speaker would deaccent "movies", although the concept "movies" seems to represent GIVEN information by any definition of the term. In (2b), it seems equally likely that both "John" and "Bill" will be accented on second mention. In (2c), the second occurrence of "president" although referring to GIVEN information, must be accented due to well-formedness constraints on prosodic structure. So, GIVEN information is not always referred to by deaccented expressions. And, in (2d), "sick" is deaccented, even though the concept of "illness" does not appear to be evoked by mention of "mother". Thus, despite the fact that information status seems intuitively to play a role in accent decisions, speakers do not always deaccent references to GIVEN entities. Similarly, speakers do not always accent references to new entities. Information status alone is not sufficient to predict speakers' decisions to accent or deaccent expressions in discourse.

Several authors have suggested further constraints that may affect accenting decisions. For example, Halliday (1967) has argued that an expression may be left unaccented if the information conveyed by the expression is anaphorically or situationally recoverable on the basis of the prior discourse or by being salient in the situation. Chafe (1974, 1976), taking a more cognitive approach, suggested that an expression may be deaccented if the information is in the listener's consciousness. It seems reasonable to assume that not all items which have been mentioned previously in a discourse of some length are recoverable anaphorically or are in the listener's consciousness, and these items may then be accented even though they represent OLD information. Others researchers have investigated speakers' decisions to accent or deaccent expressions empirically. Thus, Brown (1983) found that a more complex classification of information status predicted accent decisions fairly well: Using a taxonomy derived from Prince's (1981) classification of GIVEN—NEW information, she found that 87% of "brand new" entities (assumed not to be known by the hearer) and 79% of "new inferrable" entities (assumed by the speaker to be inferrable by reasoning from previously evoked entities) were in fact accented by subjects in a dialogue situation, while 96 — 100% of "evoked" entities (those already mentioned or situationally salient) were deaccented. Terken (1985) found a notion of "local topic" useful in helping to explain the use of referring expressions in spontaneous speech. In that study, speakers tended to use unaccented, pronominal expressions to refer to the local topic of discourse, and to use accented, full noun phrases otherwise.

These attempts to further delineate the relationship between accenting behavior and the information status of mentioned discourse entities share the intuition that the use of an

---

2 In current theories of intonational focus, (2d) might be explained by the concept of integrative accent, in which accent on a prosodic head (e.g., "mother") may convey an "all new" reading for items under the scope of that head (e.g., "sick") (Schmerling, 1976; Gussenhoven, 1983; Baart, 1987). This contrasts with "my mother got SICK" which can be used only if "mother" is GIVEN. However, speakers may also use "my MOTHER got SICK" to convey an "all new" reading, so the question as to why "sick" is accented in one case and deaccented in another remains as yet unanswered.
unaccented expression may be understood in terms of the relative ACCESSIBILITY of entities in the listener's discourse model, as estimated by the speaker. Although the notion of accessibility is not easily defined, it is intuitively clear that not all information in the listener's discourse model is equally accessible at a particular time – to serve as the antecedent of a pronoun, for instance – and that particular entities will be more likely antecedents than others. However, a drawback of the explanations given so far, especially for technological applications such as automatic text-to-speech conversion, is that there is no direct measure of accessibility. Therefore, the notion of accessibility needs to be related to observable properties of utterances in discourse (other than accentuation itself).

The relation between the relative accessibility of information in a discourse and a number of observable properties of utterances has been broadly explored in theories of "communicative dynamism" "attentional focusing" and "centering" in discourse (Chafe, 1974; Grosz, 1977; Sidner, 1979; Grosz, 1981; Sidner, 1983; Grosz, Joshi, and Weinstein, 1983; Grosz and Sidner, 1986; Kameyama, 1986; Brennan, Friedman, and Pollard, 1987; Asher and Wada, 1988; Hajicova, Kubon, and Kubon, 1990; Gordon, Grosz, and Gillion, 1993; Gundel, Hedberg, and Zacharski, 1993), and in models of sentence production (Bock and Irwin, 1980; Bock 1982, Bock and Warren 1985). The available evidence supports the notion that the relative accessibility of entities in the speaker's discourse model is a major factor in their assignment to grammatical role and surface position, and in the choice of the form of referring expressions: Highly accessible entities tend to be realized as the grammatical subject, to occur early in the utterance, and to be pronominalized. Furthermore, available evidence from studies on comprehension shows that accessibility is also an important factor in the way the listener processes the incoming message (Kameyama, 1986; Gordon et al., 1993; McKoon, Ward, Ratcliff, and Sproat, 1993). Kameyama proposes a "property-sharing constraint" governing the interpretation of pronominal expressions in adjacent utterances: If these expressions share a certain common grammatical property, e.g., if both are the sentential subject, and the first one refers to the most accessible discourse entity, the second one will be interpreted as referring to that entity also (Kameyama, 1986).

If accessibility is what really matters for deaccentuation, these observable properties should also bear a meaningful relationship to the use of deaccentuation. Indeed, Terken (1984) found a strong correlation between deaccentuation and the use of pronominal expressions. In addition, he found that deaccented expressions are associated preferably with the function of grammatical subject and tend to occur in initial position in the surface structure. Due to the limited design of the study, however, the roles of the individual factors could not be assessed independently. In the current study, we further explored the possibility that accessibility is a major factor governing the use of deaccentuation, by systematically manipulating the assignment of discourse entities to specific grammatical roles and surface positions in adjacent utterances, and by investigating the effects of these factors on speakers' decisions to accent or deaccent references to GIVEN items in a discourse.

We decided to investigate items that have been evoked by explicit mention in the immediately preceding discourse, confining our definition of GIVENness to this constrained sense. While our definition may not encompass all categories of this phenomenon, it certainly captures the core sense of the term, which all definitions of
GIVENness must include.

In order to investigate the relevance of grammatical function and surface position, and of the persistence of these properties in adjacent utterances, to deaccentuation, we compared speakers' accentuation behavior in the following scenarios: a) $X$ the referent of an expression $x$ in the target utterance, has not been mentioned in the immediate context; b) $X$ has been mentioned in the immediate context, and $x$ has the same grammatical function and/or surface position as its antecedent expression in the immediate context; and c) $X$ has been mentioned in the immediate context, but $x$ has a different grammatical function and/or surface position than its antecedent expression in the immediate context (further explanations are given below). By manipulating grammatical function and surface position semi-independently, we attempted to assess the relative effects of these factors.

To investigate the contributions of grammatical function and surface position while controlling for other factors which may also influence accent assignment, and to maintain some degree of spontaneity, we elicited productions by 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 obtained the same descriptions in different contexts, thus allowing for comparison of items with the same lexical content both within and between speakers. In this way, we overcome some problems inherent to the analysis of corpora of spontaneous speech, such as the lack of control over the context in which utterances are produced. Despite the degree of predictability this type of elicitation produces, the task exhibits the desirable property that the speaker must construct each utterance: Although the content of the utterance is dependent upon the materials displayed and the task, the actual formulation process is under the control of the speaker. The technique has been used previously in the study of prosody (Terken, 1985; Swerts and Collier, 1992).

In this way, we collected sets of utterances for a number of speakers and investigated how the assignment of accents in these productions was influenced by properties of the utterances and their prior context. In particular, we wanted to answer the following questions.

To what extent does plain GIVENness (defined as simple prior mention in the preceding context) account for the occurrence of an unaccented expression?

To what extent does the grammatical role of an expression referring to a GIVEN entity affect its accent status, and to what extent is this influence dependent on the grammatical role of the antecedent expression in the context?

To what extent does the surface position of an expression referring to a GIVEN entity affect its accent status, and to what extent is this influence dependent on the surface position of the antecedent expression in the context?
METHOD

Task

Speech was elicited from each subject in the following manner: The speaker was seated in a recording booth in front of a workstation screen displaying a simple configuration of geometrical shapes such as a circle and a triangle. The position of objects in the display could be changed in ways that had been predetermined, such that one or more objects appeared to move to different positions. The timing of such events was under the control of the experimenter. The speaker's task was to describe these events to a listener watching a second screen containing the same layout. Since the movements were few in type and very simple, they could be described in a simple, straightforward manner. By manipulating the identity and role of objects in the events and by manipulating the sequencing of events, we elicited utterances in which grammatical function, surface position, and information status (GIVEN vs. NEW) varied systematically.

Materials

Each subject described 75 "scenarios", each consisting of a sequence of four cumulative changes to an initial display ("events"); each event consisted of a change in the position of one or two objects in the visual display relative to one of the other objects in the display. By systematically varying the role of the objects involved in the events, the descriptions of these events given by our speakers could be manipulated indirectly to vary the grammatical function of the referring expressions, in the following way.

The visual display always contained nine 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 encourage lexical uniformity—each object was explicitly labelled with its name. These names remained visible throughout the experiment. Three types of events were simulated: (a) object1 could approach and touch object2, (b) object1 could approach and cover object2, and (c) object1 could approach object2 and push it against object3. So, for instance, (a) the ball could approach and touch the cone, (b) the ball could approach and cover the cone, or (c) the ball could approach the cone and push it against the cross. Objects performing particular roles in the events associate naturally with particular grammatical functions in the descriptions. For instance, the moving or pushing object is naturally assigned to the function of grammatical subject. Thus, typical descriptions of these events were: (a) "the ball touches the cone", (b) "the ball covers the cone", and (c) "the ball pushes the cone against the cross". Note that all three sample descriptions contain a grammatical subject and a direct object; description (c) also contains the object of a preposition. Only the utterance describing the fourth and last event in each scenario was used for analysis. We term this utterance the "target utterance" for our analysis, and the first three utterances in each discourse, the "context". The screen was cleared between scenarios, so that the sequence of descriptions for each individual scenario would constitute a separate discourse.

The information status of items mentioned in the target utterances was manipulated by varying the objects and the roles they played in successive events, thereby influencing the contents of successive utterances. Table 1 shows three alternative sequences of context events, a, b, and c. Each context consists of three successive changes (context events CE1—CE3) to an initial display (I) and the corresponding descriptions of the changes (CD1—CD3). Three alternative target events (TEa—c) are also presented, with
### TABLE 1

Sample context event sequences (CE) and corresponding context descriptions (CD), target events (TE) and target descriptions (TD). Any target event may follow any sequence of context events, giving nine different scenarios.

<table>
<thead>
<tr>
<th>EVENT</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Context a: constant subj</strong></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Initial Display</td>
</tr>
<tr>
<td>C Ea1</td>
<td>ball touches cone</td>
</tr>
<tr>
<td>C Ea2</td>
<td>ball touches cross</td>
</tr>
<tr>
<td>C Ea3</td>
<td>ball touches diamond</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Context b: constant obj</strong></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Initial Display</td>
</tr>
<tr>
<td>C Eb1</td>
<td>cone touches ball</td>
</tr>
<tr>
<td>C Eb2</td>
<td>cross touches ball</td>
</tr>
<tr>
<td>C Eb3</td>
<td>diamond touches ball</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Context c: constant pp-obj</strong></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Initial Display</td>
</tr>
<tr>
<td>C Ec1</td>
<td>cone pushes rectangle against ball</td>
</tr>
<tr>
<td>C Ec2</td>
<td>cross pushes line against ball</td>
</tr>
<tr>
<td>C Ec3</td>
<td>diamond pushes triangle against ball</td>
</tr>
</tbody>
</table>

**Alternative Target Types**

<table>
<thead>
<tr>
<th>TEa</th>
<th>ball touches star</th>
<th>TDa</th>
<th>the ball touches the star [subj GIVEN]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEB</td>
<td>star touches ball</td>
<td>TDb</td>
<td>the star touches the ball [direct obj GIVEN]</td>
</tr>
<tr>
<td>BEc</td>
<td>box pushes star against ball</td>
<td>TDC</td>
<td>the box pushes the star against the ball [pp-obj GIVEN]</td>
</tr>
</tbody>
</table>

---

3 In this study, the context events and their descriptions served as the non-linguistic and linguistic context, respectively, for the target event and its description, the target utterance. Since the context descriptions were accurate representations of the context events, we assume that the events and the descriptions were treated by our subjects as merged in the discourse model. Therefore, we do not distinguish in our discussion between the textual and situational context.
accompanying target descriptions (TDa—c). The target events may follow any of the three contexts. The referring expressions in the target utterances can be classified as either GIVEN or NEW, relative to the context. For instance, regardless of how we combine targets and contexts in Table 1, the item "the ball" is mentioned in all utterances of any particular context sequence, and thus will always be GIVEN (under our definition) when mentioned in the target utterance. Furthermore, the role of the GIVEN item is the same in all three context events, and the corresponding referring expression has the same grammatical function and surface position throughout the context. Thus, we assume that the speaker will consider the item "the ball" “highly accessible” in the listener’s discourse model by the time the target utterance is produced (Terken and Nooteboom, 1987). However, the other items mentioned in the target utterances (e.g., “the star” “the box”) have not been mentioned in the context and thus represent NEW information by the time the speaker produces the target utterance. The GIVEN expressions in the target utterances are the target expressions for our analysis. That is, “the ball” is the target expression in all of the target utterances in Table 1.

Orthogonal combination of targets and contexts in Table 1 produces nine different scenarios (additional scenarios were obtained by varying the identity of the objects). The scenarios differ with respect to the grammatical function of the GIVEN item in target and context. For example, the expression “the ball” has the function of grammatical subject, object, and prepositional object in contexts a, b, and c, respectively. Likewise, “the ball” fulfills these three functions in targets a, b, and c, respectively. We thus obtained a 3x3 matrix with grammatical function (subj, direct obj and pp-obj) in the context and in the target description as categorization factors. (In the following discussion we will identify these conditions in the form a → b where a gives the function in the prior context and b gives the function in the target utterance. So subj → direct obj identifies target expressions that are subjects in the prior context and direct objects in the target utterance.) In this way, we can study the effect of a change or persistence of grammatical function and the contribution of the grammatical functions involved, e.g., by pair-wise comparison of subj → subj and direct obj → subj and direct obj → direct obj and subj → direct obj. The possible contribution of properties of the current utterance to accentuation is kept constant across conditions, since each target utterance is combined with three different context types.

We also attempted to separate to some extent the effects of a change or persistence of grammatical function from the effects of change or persistence of surface position. Utterances of type TDe, containing a pp-obj, were elicited only when the pp-obj was the target expression. Similarly, utterances of type CDe were elicited only when the GIVEN item figured as pp-obj in the context, but not when the GIVEN item figured as subj or

---

4 Alternatively, speakers might have chosen to assign GIVEN items to the role of grammatical subject and/or to initial position, so that grammatical role and/or surface position would vary in the context descriptions, but this did not occur in our materials.

5 Although these items may have been mentioned in preceding scenarios, we assume that each scenario defines a separate discourse.
direct obj. Hence, in conditions direct obj \(\rightarrow\) pp-obj (context b + target c) and pp-obj \(\rightarrow\) direct obj (context c + target b), there was a change in grammatical function of the GIVEN target expression from context to target utterance, but the position in the surface structure remained the same – the last (non-empty) constituent in the utterance. By comparing these conditions with conditions in which either both grammatical function and surface position remained the same between context and target, and with conditions in which both grammatical function and surface position differed between target and context, we could separate the effect of change or persistence of surface position from the effect of change or persistence of grammatical function to some extent.

For each cell of the 3x3 design there were five scenarios, for a total of 45 target expressions elicited per subject. In addition, target descriptions a and b were combined with different contexts, in order to create two control conditions. One control was an “all new” condition, in which the expressions in the target description were not mentioned in the prior context, as illustrated in Example (3).

\[
\begin{align*}
(3) & \quad \text{CDb1: The cone touches the ball.} \\
& \quad \text{CDb2: The cross touches the ball.} \\
& \quad \text{CDb3: The diamond touches the ball.} \\
& \quad \text{TD: The rectangle touches the square.}
\end{align*}
\]

The other control was an “all given” condition, in which the subject and direct object of the last context description (CD3) were also subject and direct object in the target description, so that only the verb represented NEW information, as illustrated in Example (4).

\[
\begin{align*}
(4) & \quad \text{CDb1: The cone touches the ball.} \\
& \quad \text{CDb2: The cross touches the ball.} \\
& \quad \text{CDb3: The diamond touches the ball.} \\
& \quad \text{TD: The diamond covers the ball.}
\end{align*}
\]

No controls were included for target expressions containing a pp-obj. There were 15 scenarios for each control condition.

Subjects

Ten male native speakers of standard American English, employed at AT&T Bell Laboratories in Murray Hill, New Jersey, participated as volunteers.

Procedure

Each subject was treated individually and received oral and written instructions. He was informed that he was participating in an experiment aimed at improving the way computers provide spoken information to humans. He was told about the basic structure of the experiment, about the possible events he would see in the display, and about the
format of the scenarios. He was then given examples of possible descriptions, in order to inform him about the degree of precision needed in the descriptions. He was informed that there would be someone listening to him to verify his descriptions relative to the events in another display, and was introduced to the listener (J. Hirschberg), but he was told that there would be no interaction or feedback, either auditory or visual. He was told that the best speaking mode would be as if telling a story or describing a movie he had seen, to the listener.

During the experiment, the subject and the experimenter were in a soundproof booth. The subject was seated before the screen of a SUN 361 workstation. The actual events and their order were predetermined and read from a file. The pacing of events was controlled by the experimenter: Following the speaker's description of an event in the display, the experimenter pressed the return key of the workstation keyboard, triggering the display of the next event. The events all involved a change of the relevant object's or objects' position in a number of steps to a new position. Between scenarios, the screen was cleared. The actual experiment was preceded by ten practice scenarios, to familiarize the subject with the task. Upon request, the experimenter made clear that object color and size and orientation of an object relative to another object were not relevant. No subject reported difficulty with the task.

The actual experiment took about 20 minutes for each subject. Halfway through the scenarios there was a short break. Some subjects tended to speak in a very low voice, possibly due to the artificiality of the communicative situation or initial hesitancy about the recording situation. In such cases, they were encouraged to speak up a bit and they were reminded that they were actually speaking to a listener.

The scenarios were presented in 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. The descriptions were stored on DAT (sampling frequency 48 kHz).

Prosodic analysis

The descriptions obtained had the same format as those in Table 1: simple active sentences of the form "the NOUN VERBs the NOUN" or "the NOUN VERBs the NOUN PREP the NOUN" Pronouns and passive constructions, which are commonly used by speakers to establish discourse coherence, did not occur in our materials. We deliberately discouraged the use of pronominal expressions by explicitly labelling objects to be described in the experiment with their names. One might object that this artificiality reduces the usefulness of the results as data relevant to the question of why speakers use deaccentuation to mark discourse coherence. However, it was necessary to limit the number of factors involved in the experiment in order to study the relationships among a manageable few, which could be manipulated in a principled way.

Each subject's utterances were stored on disk in separate files and downsampled to 12 kHz. The target utterances were then excised from the context for further analysis. All target utterances were analysed by ear by the two authors independently, to determine the location of pitch accents. The two judges had no information about the context of a target description, the condition to which a target item belonged, or the judgment of the other judge. For each referring expression in a target utterance each judge indicated whether
the expression was accented or not on a three-point scale: "unaccented" (0), "uncertain whether accented or unaccented" (1), and "accented" (2). For each individual expression an "accentedness score" between 0 and 4 was obtained by summing the numerical values assigned by the judges. This procedure combines accentedness judgments and confidence ratings, and appears to be relatively free of theoretical and response biases.

Table 2 shows the data for agreement between judges on the accentedness of referring expressions in target utterances.

The marginals in Table 2 show that both judges judged about the same percentage of expressions to be accented; the two judges differed, however, in the percentage of expressions they judged deaccented. These facts suggest that the judges' thresholds for deciding that an expression was accented were about the same, but that judge A was more cautious than judge B in classifying an expression as unaccented. Next, consider the percentages for full agreement (summing cells 00, 11 and 22), moderate disagreement (01, 10, 12 and 21) and full disagreement (02 and 20); the percentages for these are 81.3%, 11.6% and 7.1%, respectively (barring small rounding errors). Subtracting the entry in cell 11 (shared uncertainty) from the percentage full agreement gives 80.8% agreement about the absence or presence of an accent (cells 00 and 22), which appears acceptable. The agreement data for each individual condition (\textit{subj $\rightarrow$ subj}, \textit{subj $\rightarrow$ direct obj}, \textit{direct obj $\rightarrow$ subj}, ...) were similar. This means that the overall results are not due to unequal distributions of clear and less clear cases of deaccentuation in the different conditions (although there seemed to be less agreement for grammatical subjects than for direct objects and pp-objects).
TABLE 3

Mean accentedness scores as a function of grammatical role and surface position in the target utterance and in context. Part (a) gives data for target expressions, (b) for non-target expressions in target utterances, and (c) for controls. For (a), each cell represents 50 observations, for (b) and (c) each cell represents 150 observations.

<table>
<thead>
<tr>
<th>Grammatical Role / Surface Position</th>
<th>in Context</th>
<th>in Target Utterance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>subj</td>
<td>direct obj</td>
</tr>
<tr>
<td>(a) GIVEN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subj</td>
<td>2.1</td>
<td>3.6</td>
</tr>
<tr>
<td>direct obj</td>
<td>3.3</td>
<td>0.6</td>
</tr>
<tr>
<td>pp-obj</td>
<td>3.0</td>
<td>1.4</td>
</tr>
<tr>
<td>(b) NEW</td>
<td></td>
<td>3.7</td>
</tr>
<tr>
<td>(c) Controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>all given</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subj</td>
<td>0.6</td>
<td>–</td>
</tr>
<tr>
<td>direct obj</td>
<td>–</td>
<td>0.2</td>
</tr>
<tr>
<td>all new</td>
<td>–</td>
<td>3.3</td>
</tr>
</tbody>
</table>

RESULTS AND DISCUSSION

The main results are presented in Table 3. The cell entries represent mean accentedness scores, averaging across replications and speakers. Table 3(a) shows accentedness scores for GIVEN target expressions as a function of grammatical role and surface position in the context and in the target utterance. Table 3(b) shows scores for non-target noun phrases in target utterances, i.e., expressions referring to items which

---

6 A mean accentedness score of 4.0 would indicate that both judges categorized all expressions in a particular cell as accented, a mean score of 0.0 would mean that they categorized all expressions in a cell as unaccented. Since agreement between judges about the presence or absence of an accent was high and did not vary strongly across conditions, a mean accentedness score may be interpreted loosely as a +accent percentage. For example, mean accentedness scores of 0.0, 1.0, 2.0, 3.0 and 4.0 would correspond roughly to 0, 20, 40, 60, 80 and 100% +accent expressions.
had not been mentioned in the preceding context utterances and are NEW in the current scenario. Note that Table 3(b) contains data for subj and direct obj only, since target descriptions containing pp-objs were used only with GIVEN pp-objs. Table 3(c) contains data for two control conditions. Target utterances in these conditions also contain only subj and direct obj expressions. In the “all given” control, both expressions were GIVEN and shared grammatical function and surface position with the previous utterance (cf. Example (4) above). In the “all new” control, both expressions were NEW, i.e., not mentioned in the context utterances of the scenario (cf. Example (3) above). Note from Table 3(c) that in the “all new” control, both the subject and the object tend to be accented. This suggests that speakers treated subject and object as separate focal domains, each with its own prosodic head. It therefore seems appropriate to treat an unaccented subject in this experiment as a case of true deaccentuation – rather than just as an “unaccented expression” within the (metrically defined) scope of an accent on the direct object (Gussenhoven, 1983; Dirksen, 1992).

The important results for our study appear in Table 3(a). It can be seen that average accentedness scores for target expressions vary considerably across context conditions. To evaluate these differences, mean accentedness scores were computed for each condition for each speaker (collapsing over replications), and these data were subjected to an analysis of variance with grammatical function in the prior context and the target utterance, respectively, as fixed factors, and speakers as replications. An analysis with items as the replication factor was not performed in view of the homogeneity of the items. The two main effects and their interaction were all found to be significant [F (2,18) = 13.87, p < 0.0003, F (2,18) = 21.97, p < 0.0001, and F(4,36) = 28.5, p < 0.0001, respectively].

Due to the strong interaction it makes little sense to make general statements about the contributions of grammatical function and surface position in the current utterance as if, without taking into consideration the grammatical function and surface position of the antecedent expression in the context. The significant interaction seems to have several components. First, we note that the values on the diagonal in Table 3(a) are lower than the off-diagonal values. This means that expressions for which the grammatical function in the current utterance and in the context are the same, i.e., where there is persistence of grammatical function, are more likely to remain unaccented than expressions for which there is a change in grammatical function. Thus, persistence of grammatical function appears to be a relevant factor in triggering deaccentuation. It may also be noted that some of the off-diagonal values in Table 3(a) are as high as those for the “all new” conditions in Table 3(c). This is consistent with the observation that simple GIVENness in the sense of “previously mentioned in the context” does not itself suffice to trigger deaccentuation.

Second, the values on the diagonal in Table 3(a) seem to fall into two classes: subj → subj has a higher accentedness score (2.1) than the other two conditions, direct obj → direct obj and pp-obj → pp-obj (0.6 and 0.7, respectively). This is surprising, since it has been argued that subjects are more likely to be deaccented than other expressions (Nootbeoom and Terken, 1982). Our findings here may reflect a tendency towards rhythmic alternation of accents in pre-nuclear position, i.e., before the final accent in a prosodic phrase (Horne, 1990). Comparison of the values for subj → subj in Table 3(a) with those for subj → subj in the “all given” condition in Table 3(c) lends support to this interpretation. While the subj → subj condition of Table 3(a) represents
sequences such as (5), the subjects in the “all given” condition of Table 3(c) represent sequences such as (6) (cf. example (4) above).

(5)   a  The ball touches the CONE.
      b  The ball touches the BOX.

(6)   a  The ball touches the CONE.
      b  The ball COVERS the cone.

For pragmatic reasons, utterance 5(b) has an accented object noun, and utterance 6(b) has an accented verb. If we represent unaccented content words by \( w \) for *weak* and accented content words by \( s \) for *strong*, the metrical structure for 5(b) would be \( wws \), but for 6(b) it would be \( wsw \). The rhythmic alternation hypothesis holds that there is a tendency to alternate strong and weak nodes in a prosodic phrase, thus converting \( wws \) into \( sws \); the result of this conversion would be the accenting of “ball” in 5(b). But since 6(b) already exhibits rhythmic alternation, no conversion would occur, and “ball” would remain weak and unaccented. The absence of a difference between the accentness scores for \( \text{direct obj} \rightarrow \text{direct obj} \) in Table 3(a) and the “all given” condition in Table 3(c) suggests that rhythmic alternation applies only before the last accent in the phrase.

An alternative explanation for the finding that subjects tend to be accented more often than either direct objects or pp-objects is that, in pre-nuclear position, the distinction between GIVEN and NEW is not signalled by the absence or presence of an accent but by varying the pitch range, i.e., by creating a difference in the phonetic realization of the accents (Kruyt, 1985; Horne, 1991). In order to test this possibility, we computed average \( F_0 \) maxima for accented GIVEN and NEW grammatical subjects. To avoid contamination of the data by judges' uncertainty, we included only those cases where judges agreed that an expression was accented. To avoid confounding with vowel intrinsic \( F_0 \) contributions, we included only those cases where tokens of the same utterance with an accented target expression were available in all six relevant conditions: \( \text{subj} \rightarrow \text{subj}, \text{direct obj} \rightarrow \text{subj} \) and \( \text{pp-obj} \rightarrow \text{subj} \) for the GIVEN conditions; \( \text{subjects} \) following contexts with constant \( \text{subjects}, \text{objects}, \) and \( \text{pp-obj} \), respectively, for “all new” utterances. In all, there were only nine sets of utterances which had accented target expressions in all six relevant conditions, from six different speakers.

\( F_0 \) measurements for these utterances were obtained using a cross correlation algorithm (Talkin, 1989). \( F_0 \) maxima were measured at the amplitude peak of the stressed syllable of the subject noun. Average \( F_0 \) maxima for the different conditions were 123.3 Hz for \( \text{subj} \rightarrow \text{subj} \), 119.1 for \( \text{obj} \rightarrow \text{subj} \), 119.9 for \( \text{pp-obj} \rightarrow \text{subj} \) and 120.8, 120.1 and 121.0 for the subjects in the different “all new” conditions. These results do not confirm Kruyt's and Horne's findings; there is very little difference in our data between accented GIVEN and NEW subjects. Instead, these data suggest that speakers made a simple binary decision to accent or to deaccent an expression. Once that decision was made, subjects realized it acoustically in very predictable ways. However, since the number of data points from which the means are computed is very small, this conclusion should be treated with some caution.
Third, the off-diagonal values for direct obj $\rightarrow$ pp-obj and pp-obj $\rightarrow$ direct obj (1.6 and 1.4, respectively) are lower than the other off-diagonal values (3.6, 3.2, 3.3 and 3.0). Note that in the present stimuli direct obj $\rightarrow$ pp-obj and pp-obj $\rightarrow$ direct obj had different grammatical functions but the same surface position in context and target, while the other conditions had both different grammatical functions and different surface positions in the context and the target utterance. So, the differences in mean accentedness scores appear to be due to persistence vs. change of surface position. At the same time, the scores for conditions with different grammatical functions but the same surface position in context and target (1.6 for direct obj $\rightarrow$ pp-obj and 1.4 for pp-obj $\rightarrow$ direct obj) are higher than those for conditions with both the same grammatical function and surface position in context and target (0.6 for direct obj $\rightarrow$ direct obj and 0.7 for pp-obj $\rightarrow$ pp-obj). This means that both persistence of grammatical function and persistence of surface position contribute to deaccentuation, although there appeared to be differences between speakers with respect to the relative weight of persistence of grammatical function and surface position.

**Speaker variability**

In order to explore whether the bimodal distributions of accentedness scores observed in some conditions reflect different speaker strategies or just random variation within speakers, mean scores for each speaker were computed for the following situations: (a) target expressions with the same grammatical function and surface position in context and target utterance (direct obj $\rightarrow$ direct obj and pp-obj $\rightarrow$ pp-obj), (b) same surface position but different grammatical functions in context and target utterance (pp-obj $\rightarrow$ direct obj and direct obj $\rightarrow$ pp-obj), and (c) a shift in grammatical function and surface position between context and target utterance (subj $\rightarrow$ direct obj and subj $\rightarrow$ pp-obj). Due to the setup, the condition where target expressions had the same grammatical function but different surface positions in context and target utterance was missing. The scores are shown in Table 4.

These data suggest that speakers fall into three subgroups with respect to their accenting strategy. For two speakers (3 and 4), conditions (b) and (c) have about the same average accentedness score and condition (a) has a lower score. For these speakers, “same function” appears to be a necessary condition for deaccentuation. For five speakers (1, 2, 6, 7 and 10), there is a large difference between conditions (a) and (b), on the one hand, and condition (c), on the other. For these speakers, “same position” appears to be a sufficient condition for deaccentuation. Three speakers (5, 8 and 9) appear to differentiate between all three conditions. For these speakers, “same position” and “same function” appear to be more or less independent conditions, and the more the conditions influencing deaccentuation are satisfied, the more deaccentuation occurs.

---

\[ F(1,9) = 9.7, p < 0.01 \] but not for \[ F(1,9) = 2.8, p = 0.13 \].
TABLE 4

Mean accentedness scores for target expressions for individual speakers as a function of persistence of grammatical function and/or surface position

<table>
<thead>
<tr>
<th>Condition</th>
<th>(a) same position same function</th>
<th>(b) same position different function</th>
<th>(c) different position different function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 0.3</td>
<td>0.4</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>2 1.1</td>
<td>1.2</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>3 0.0</td>
<td>3.2</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>4 1.8</td>
<td>2.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Speaker</td>
<td>5 1.1</td>
<td>1.8</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>6 1.0</td>
<td>1.5</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>7 0.0</td>
<td>0.4</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>8 0.5</td>
<td>2.0</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>9 0.0</td>
<td>1.0</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>10 0.5</td>
<td>1.0</td>
<td>4.0</td>
</tr>
</tbody>
</table>

GENERAL DISCUSSION

Our findings show that simple GIVENness, operationally defined as mere mention of an item in the immediate context, is not a sufficient condition for deaccentuation. Instead, additional factors, the persistence of surface position and the persistence of grammatical function from the context to the current utterance, play an important role. If an expression has the same grammatical role and surface position as its antecedent expression in the immediate context ("persistence of grammatical function and surface position"), it is likely to be deaccented. If an expression has a different grammatical function than the antecedent expression but occupies the same surface position ("persistence of surface position"), it is also likely to be deaccented, but less likely than in the former condition; in addition, the results suggest that speakers may employ different strategies in this situation. If there is a change in both grammatical function and surface position from one utterance to the next, a GIVEN expression is in fact likely to be accented – as likely as a NEW expression.

The data on the accentuation of grammatical subjects lend some support to the hypothesis that the occurrence of deaccentuation may be affected by the rhythmical
properties of the utterance. Our speakers’ accentuation of these expressions might be explained as the application of some form of rhythmic alternation of accented and unaccented words before the final accent in a prosodic phrase.

What answers do these results provide to the question of why expressions referring to GIVEN information are accented in some cases and unaccented in others? If we assume that the distribution of unaccented expressions can be used to draw inferences about the function of deaccentuation for the listener, our findings suggest that grammatical function and surface position may serve as cues to the mapping of referring expressions onto entities in the mental representation of the preceding discourse. The grammatical and positional characteristics of unaccented expressions may be used as cues to access candidate antecedents in the discourse model. We may speculate that such antecedents are made accessible from the current expression by virtue of this property-sharing. When an expression does not share these properties with its intended antecedent, but with a non-antecedent expression or expressions, the intended antecedent will be less readily accessible, due to interference from these competing expressions. We assume that the reduced accessibility may be signalled by the accenting of the target expression. In this way, deaccentuation contributes to the pruning of the set of candidate antecedents.

This interpretation suggests that accessibility is indeed what matters in deaccentuation, but not just the relative accessibility of entities in the discourse model, without taking into consideration the properties of the referring expression itself. Rather, accessibility is to be defined in terms of how easily an entity in the discourse model can be accessed from the referring expression, on the basis of the properties of both the referring expression and the entity in the discourse model.

Clearly, one would not want to argue at this stage of research that syntactic and surface position information represent the only relevant sources of information in the mapping of expressions onto entities in the discourse model. Other sources of information such as lexical and morphological information, are obviously important for antecedent identification, not only in cases of lexical repetition and morphological relationship, but also in matters of synonymy and gender/number information. The relative contribution of each of these factors remains to be explored.

Our findings are relevant to current cognitive approaches to the mapping process. For instance, centering theory, which has been applied primarily to the solution of pronominal expressions (Grosz, 1981, Grosz et al., 1983; Gordon et al., 1993), defines for each non-initial utterance \( u \) in a discourse segment a “backward-looking center” (\( C_b \)) which serves to link the utterance to the previous discourse, and a set of “forward-looking centers” (\( C_f \)), including \( C_b \), which provide potential links to the subsequent utterances in the discourse. After utterance \( u \) has been uttered and integrated into the speaker’s and listener’s discourse model, the elements in \( C_f \) differ with respect to their prominence in the discourse model. The prominence ranking is affected by properties such as grammatical role, surface order, surface realization of the reference (e.g., pronoun vs. full noun phrase), and the accentuation of the expressions. In this way, centering theory captures the notion that items in the discourse model may have different degrees of accessibility. For the speaker, the prominence ranking of entities in the discourse model affects the way he may refer to them in the subsequent utterance. For the listener, the prominence ranking affects the interpretation of the referring expressions contained in the subsequent utterance: The
prominence ranking provides an ordering for evaluating candidate interpretations of pronouns in the next utterance. Current formulations of centering theory seem to assume that surface position and grammatical function of expressions are not directly involved in the mapping process, although Kameyama (1986) has proposed to incorporate a "property-sharing" constraint.

On the other hand, McKoon et al. (1993) assume that accessibility is a function of the properties of both entities in the discourse model and the current expression. The present findings provide further support for this assumption, for the following reason. The experimental scenarios in our study were constructed so that the context-setting utterances preceding the target would favor a single entity $X$ as the most prominent item in the set $C_{i-1}$ at the onset of the target utterance $C_{i+1}$. In the absence of factors enhancing the prominence of other entities in the discourse model, one might expect that, at the onset of the target utterance, $X$ should have higher prominence than any competing entity. If the prominence of $X$ in the set $C_{i-1}$ were the only property of $X$ relevant to the mapping process, one would expect deaccentuation to occur in all cases in which the target utterance $C_{i+1}$ in fact contained an expression $x$ referring to $X$. However, in our data this clearly did not occur. Instead, deaccentuation occurred most frequently when $x$ in utterance $C_{i+1}$ shared grammatical function and/or surface position with its antecedent expression in utterance $C_i$. Unless we assume that these properties are represented explicitly in the representation of elements of $C_{i-1}$ and are used directly in the mapping process, it is difficult to explain the pattern of results.

A possible limitation to this conclusion might be that the strong effects of persistence of grammatical function and surface position may arise only due to the syntactic parallelism of successive utterances in our context and target utterances. In situations where there is more syntactic variation between consecutive utterances, these factors might constitute less useful cues to the process of establishing a link between a referring expression and its antecedent, compared to other sources of information. For instance, in less constrained situations speakers employ other kinds of devices such as topicalization or passivization to maintain coherence, which may result in syntactic variation. The possibility that persistence of grammatical function and/or surface position is less relevant to deaccentuation in such situations should be investigated in future research.

It might also be suggested that our speakers produced a high percentage of accented expressions for GIVEN items because our experiment inadvertently fostered a "contrastive" situation, in which speakers were always contrasting one geometric shape with other members of a small set of such shapes and thus accented their expressions to indicate contrast. The possibility of reinterpreting our findings in this light does not seem useful in explaining the differences in accentuation behavior we observed, since the possibilities for using contrastive stress were available for all speakers in all conditions of the experiment. Alternatively, one might suggest that the contrast was not between a specific object and all other elements of the set of objects used in the experiment, but between a specific object and the one that was mentioned in the previous utterance with the same grammatical role and/or in the same position. However, this interpretation is close to our own interpretation, as it appears to imply the same assumptions about the mechanisms involved in mapping a referring expression onto possible antecedents in the discourse model.
To summarize, we started from the observation that GIVENness by itself is not a sufficient condition for deaccentuation. We found that persistence of grammatical function and persistence of surface position are additional constraints on deaccentuation. We hypothesize that speakers use deaccentuation when they believe that an intended interpretation is easily accessible in the discourse model. Therefore, we propose that the properties of grammatical function and surface position may be used by the listener as cues to access potential antecedents in the discourse model, by leading him or her to look for a candidate antecedent with the same properties. However, we leave open the possibility that the cue validity of grammatical function and surface position in relation to other sources of information may vary with changes in the characteristics of the communicative situation. This hypothesis remains to be tested in further investigations.

(Received February 7, 1994; accepted March 17, 1994)

REFERENCES


