

# Informational Status and Pitch Accent Distribution in Spontaneous Dialogues in English

K. Ito, S. R. Speer & M. E. Beckman

Department of Linguistics  
Ohio State University, USA

{ito;speer;mbeckman}@ling.ohio-state.edu

## Abstract

Revealing the relations between pitch accent types and the informational status of words requires a refined discourse analysis of spontaneous speech. A cooperative unscripted task in which subjects gave instructions for decorating Christmas trees successfully induced production of target adjective-noun pairs conveying new/given and contrastive information. Adapting Grosz and Sidner's intention-based discourse analysis [1], each target word was tagged for its newness or givenness and also for *contrastiveness* at both the *discourse* level and the *discourse segment* level. The analyses show that contrastiveness was a good predictor of accent type (L+H\*), and that the finer-grained discourse segment level analysis was somewhat better than the discourse level in predicting the presence or absence of accent. Local word position (adjective or noun) interacted with both contrastiveness and discourse segmentation in the assignment of accent.

## 1. Introduction

A speaker's production of pitch accents in an utterance may be driven by multiple communicative motives, including conveying the speaker's emotional status, marking the beginning of rhythmic units that allows a robust parse of the prosodic structure, and emphasizing the importance of specific information in the message. One of the most-investigated functions of accentuation is to mark of the new vs. given status of a word in a discourse. It has been claimed that in languages such as English, German and Dutch, speakers produce pitch accents on words that are newly introduced to a conversation, whereas words carrying given or already-mentioned information tend to be produced without accents [2, 3, 4]. However, more recent studies provide counter-evidence to the idea of such a simple one-to-one correspondence between the presence vs. absence of an accent and the new vs. given status of a word [5, 6]. For example, studies of English have shown that a word may be accented when it is re-introduced in the discourse [5]. In addition, a previously-mentioned word may be produced with a particular type of rising pitch accent when it carries contrastive information in relation to other words or discourse elements (e.g. the words *Spanish* and *spinach* are contrastive to each other in *I ordered a SPANISH omelet, not a SPINACH omelet!*). This "contrastive" rising accent has different tone scaling and alignment from accents that mark novel information [7, 8]. While the use of accentuation to mark newness has been studied extensively, the assignment and distribution of different pitch accent types has been investigated only with carefully scripted productions. This paper introduces a technique for eliciting spontaneous dialogues which can be used to examine the role of

contrastiveness and its interaction with discourse segmentation and the new/given distinction. The experiments used an unscripted cooperative task where a pair of speakers decorated a series of Christmas trees. Ornaments to be hung on the trees were indicated by labeled photographs presented on a computer screen. Labels had varying combinations of color-term adjectives and object nouns (e.g., green ball). The order of ornament presentation was varied in order to manipulate the new/given and contrastive statuses of these adjectives and nouns.

Even with an extended recording of spontaneous speech, examination of the intonational characteristics of the target words would be misleading if the informational statuses of those words are misidentified. In this study, a model of intention-based discourse analysis [1, 6] is employed to give a finer-grained definition of newness/givenness of the target words. Subjects' utterances mentioning the target ornaments are grouped according to the speakers' local communicative goals or *discourse segment purposes*. The new/given status of a word is determined at two levels by examining whether it has been mentioned in the discourse (i.e., the current tree) or within the current discourse segment (DS) (e.g., a row of ornaments or a local sequence of ornaments). The analysis of the utterances of two speakers indicated that deaccentuation was more tightly linked to givenness at the DS-level level than givenness at the D-level. In addition, target words were labeled for their contrastive status (e.g. *To the right of the blue ball, please hang an ORANGE ball.*: orange=contrastive, ball= non-contrastive). An analysis of accent type showed that a rising accent (L+H\*) was most likely to be produced in a contrastive position, while its use was rare in non-contrastive positions. Finally, the position of the word (i.e., adjective vs. noun) was found to be a major factor in predicting the location of L+H\* accents and the occurrence of *deaccentuation*.

## 2. Experiment

### 2.1. Design and Materials

Participants were told that they would work in pairs to decorate Christmas trees. A confederate teamed up with each participant, and had the participant work in the role of the 'director' to give instructions as to what ornament to pick and where to place it. Each participant gave instructions to complete 4 trees, each with 24 target and 8 distracter ornaments. Each ornament was described using a combination of a color adjective (targets: navy, green, orange, gray, beige, blue, brown; distracters: purple, white, gold, silver) and an object noun (targets: ball, egg, onion, house, bell, drum, doll, candy; distracters: snowman, stocking, star, hat). Within a tree, the target adjectives and nouns appeared three times each. Identical adjective-noun pairs were never repeated

within a tree, and the arrangement of the ornaments was controlled to elicit an equal number of first mentions vs. second/third mentions. Givenness of an item name was induced by either consecutive mention of color adjective or object noun (e.g. blue in the sequence *blue ball – blue house*) or distant mention, i.e. the word was repeated after several intervening ornament trials. Each target adjective or noun appeared in a consecutive trio once within the four trees, so that each tree had 2 color and 2 object trios (e.g. adjective trio: *green candy - green ball - green bell*; object trio: *orange house – brown house – gray house*).

## 2.2. Participants

16 undergraduate students at Ohio State University, who were native Midwestern American English speakers, participated in the experiment as a part of their linguistics course requirement. The confederate was an undergraduate assistant, who acted as if he were participating in the experiment for the first time in each session.

## 2.3. Procedure

Participants (*directors*) were seated in front of a computer monitor in a soundproofed booth, and wore headphones and a microphone to communicate with the confederate *decorator*, who was seated outside the booth. Through a window in the booth wall, the directors could see the Christmas tree and the hands of the decorator who followed their instructions. Directors could see neither the face of the decorator nor the ornament tray in front of him. On each trial, an animated display presented a photo of the next ornament on the right side of the screen, and a tag naming the ornament (e.g. *orange drum*) appeared at the location on the tree where it should hang on the left side of the screen (see Figure 1). The director told the decorator what ornament to choose and where to place it according to the display. The timing of the display was controlled by the experimenter, who sat outside the booth.



Figure 1: Example Christmas tree display.

Directors were not explicitly instructed about what to say, nor were they told to use the tags to name ornaments (but most of the participants used the tags). The decorator picked the ornament and showed it to the decorator through the booth window, asked whether the correct ornament was chosen, and hung it according to the instruction. After the director confirmed through the window that the ornament was hung correctly, the decorator asked for the next instruction and the

experimenter advanced the display for the next trial. The confederate decorator was instructed not to repeat the target words but to use expressions such as ‘this/that (one)’ while interacting with the director, so that the decorator’s utterances would not affect the new/given status of the target words for the director.

## 2.4. Analysis of Informational Status

Digital recordings of the sessions were transcribed including both directors’ and decorator’s speech. Each target word (spoken only by directors) was tagged for its new/given status at two different levels: Discourse-new/given (D-n/g) and Discourse Segment-new/given (DS-n/g). An example transcript with two Discourse segments is shown in Table 1. Each tree served as the largest unit of discourse for the analysis. A word was considered Discourse-new (D-n) when it was mentioned for the first time in the currently decorated tree. The second, third, and any subsequent mentions within the same tree were marked as Discourse-given (D-g). A *Discourse Segment* (DS) was defined as ‘a block of utterances exchanged between the two participants to achieve a local communicative goal such as completing subsection of the tree.’

Table 1: Example text transcription of dialogue.

54.2	Dir: At the very top there's a white hat.
	Dec: [shows] That one?
	Dir: Yeah.
	Dec: All right. [places] Okay. Next?
069.5	Dir: Uh, a blue (H*) house (H*).
	[Dec: shows through booth window]
	Dir: Yeah.
	Dec: And which way did you say ... right or left?
	-----Discourse Segment Boundary-----
	Dir: Uh, it's gonna start on the left.
	Dec: [places on tree] There we go. Okay. Next.
095.3	Dir: Uh, blue (L+H*) bell (H*).
	[Dec: shows through window]
	Dir: Yeah.
	Dec: [places on tree] Okay.
108.9	Dir: Uh, this is an orange (L+H*) bell ( ).
	-----Discourse Segment Boundary-----
	[Dec: starts to put on tree]
	Dir: And yeah, there'll be... my bad, I forgot to tell you, this is ... this is gonna be on the next down; there'll be three in this row.
	-----Discourse Segment Boundary-----

When a word was mentioned for the first time in such an utterance block, it was tagged as Discourse Segment-new (DS-n), while if the word appeared more than once within a DS, its non-initial mentions were tagged as DS-g. Thus, a word that appeared for the first time in a given tree was marked as D-n and DS-n, whereas a word that had been mentioned in an earlier trial was tagged as D-g, but also as DS-n if it appeared in a new DS. When a word was repeated within the same DS, its status was marked as both D-g and DS-g (see Table 2). In addition to the above new/given marking, target words were also analyzed for *contrastiveness* in each word position. Contrastiveness was defined according to the presence of an immediately preceding reference to an ornament described with the same noun but different adjective or with the same adjective but different noun (e.g., *blue* is contrastive in a sequence *green ball-blue ball*, so is *drum* in *brown house-brown drum*). As with new/given, contrastive

status was defined either locally (i.e., the preceding reference was in the same DS) or globally (i.e., a DS boundary intervened between the two references). It was predicted that a word bearing contrastive status would be produced with L+H\* more often than words with non-contrastive status.

Table 2: Summary of the new/given status tagging.

WORD POSITION	NEW/GIVEN	
	D-level	DS-level
Adj/Noun	D-n	DS-n
Adj/Noun	D-g	DS-n
Adj/Noun	D-g	DS-g

### 2.5. ToBI Annotation

Each utterance that included a target adjective and/or noun was submitted to ToBI transcription [9] by two experienced transcribers. Praat [10] was used for the display of spectrograms and F0 traces that were necessary for the identification of word boundaries and tonal shapes.

## 3. Results and discussion

Target utterances produced by two speakers were analyzed closely with the tagging scheme described above. The frequency of accentuation was calculated for each word position in each of the four combinations of new/given status between the adjective and the noun: NewNew, NewGiven, GivenNew, GivenGiven. All accentual types except for L\*, which is used to query rather than indicate the insertion of information into the common belief space [8], were counted as cases of accentuation on a word.

### 3.1. Distribution of Accents

Table 3 summarizes the counts of the accentuation in each new/given condition at the D-level and the DS-level for the two speakers. Our results indicate that the simple definition of new/given status as a distinction between previously-mentioned vs. not-previously-mentioned words does not predict the presence or absence of accentuation in spontaneous speech. First of all, word position was an important factor for accentuation. On one hand, target adjectives were produced with an accent more than 80% of the time regardless of their new/given status at both D- and DS- levels. Even when the word was repeated within a local discourse segment (in GN and GG conditions), an adjective bore an accent 85% of the time. On the other hand, nouns tended to be produced without an accent more frequently when they had *given* status, at both D- and DS-levels. *New* nouns bore an accent only slightly less often than adjectives, on 83% of trials, but *given* nouns bore an accent on 58% of trials.

The new/given status of the preceding adjective also affected the accentual distribution on the nouns. The words least likely to bear an accent were *given* nouns, and those preceded by *new* adjectives were less likely to bear accent than those preceded by *given* adjectives. This was the case even though many of the nouns were in phrase- or utterance-final position, where under other conditions a noun would be likely to bear a nuclear accent. We refer to these nouns as *deaccented*. *Given* nouns were more likely to be deaccented

when preceded by a *new* than a *given* adjective. This was true for NG vs. GG at both D and DS- levels. Note, however, that even in the NG condition at DS level, the target nouns were accented 39% of the time. (Since there were only six tokens of GG at the DS-level, no generalization should be drawn about the frequency of accentuation for these *given* words until sufficient observations are obtained from a larger group of subjects.)

Table 3: Proportion of accentuation and number of observations for target adjectives and nouns in each new/given condition.

DOMAIN	NEW/GIVEN	# OF TRIALS	ADJECTIVE	NOUN
Discourse (D)	NN	31	0.94	0.84
	NG	32	0.84	0.50
	GN	32	0.91	0.81
	GG	96	0.89	0.67
Discourse Segment (DS)	NN	67	0.91	0.88
	NG	62	0.89	0.39
	GN	56	0.88	0.79
	GG	6	0.83	0.83

Although the words marked as *new* at the DS level were not necessarily accented more frequently than the words marked as *new* at D-level, the *given* words at the DS-level were apparently deaccented more frequently than the *given* words at the D-level. This may suggest that the discourse segment level of structure plays a more important role in constraining the *deaccentuation* of given words than in licensing accentuation on new words.

### 3.2. Contrastiveness and Accent Type

Table 4 summarizes the counts of L+H\* accents in each new/given condition at the D-level and the DS-level for the two speakers. As the table shows, these accents were far more likely to appear on both adjective and noun target words in contrastive contexts than on comparable words in non-contrastive contexts, and this was so at both the D- and DS-levels. The word position effect on accent assignment seen for accents in the previous analysis was also found for L+H\* accents, so that an adjective in a contrastive context was more likely to bear an accent than a noun in a contrastive context. Both *new* and *given* adjectives in contrastive contexts bore a L+H\* accent on about half of the trials, while *new* and *given* nouns in contrastive contexts at the D-level, and *new* nouns in contrastive contexts at the DS-level bore a L+H\* accent on only about one-fifth of the trials. Although no *given* nouns in contrastive contexts at the DS-level were observed to have a L+H\* accent, there were only two such trials. Thus no generalization should be drawn about accent type for these *given* words until sufficient observations are obtained from a larger group of subjects.

Table 4: *Proportion of L+H\* accents and number of Contrastive and Non-Contrastive context trials for target adjectives and nouns in each new/given condition under two discourse tagging systems.*

DOMAIN	N/G	STATUS	ADJECTIVE		NOUN	
			# OF TRIALS	L+H*	# OF TRIALS	L+H*
Discourse (D)	NEW	Contrastive	26	0.46	31	0.19
		Non-Contrastive	37	0.03	33	0.06
	GIVEN	Contrastive	37	0.51	33	0.18
		Non-Contrastive	91	0.00	95	0.00
Discourse Segment (DS)	NEW	Contrastive	60	0.47	56	0.18
		Non-Contrastive	101	0.04	67	0.04
	GIVEN	Contrastive	3	0.67	2	0.00
		Non-Contrastive	59	0.03	69	0.04

As suggested by the above results, the same ornament names were produced with quite distinctive tonal patterns depending on the words' informational status. Figure 2 shows the example F0 contours of *green candy* produced in two different contexts: (a) preceded by *green house* within the same DS (left) vs. (b) preceded by *beige candy* in the same DS (right). Thus, the contrastive information was conveyed by the noun *candy* in the former utterance but by the adjective *green* in the latter. When a given adjective *green* preceded a contrastive noun *candy*, it was still produced with a clear pitch accent H\*. The contrastive adjective *green* was produced with a L+H\*, which was followed by a very compressed pitch range for the noun. No accent was identified for the noun after the contrastive L+H\*. When the noun was conveying contrastive information, it was produced with a downstepped rising pitch accent !H\* instead of L+H\*.

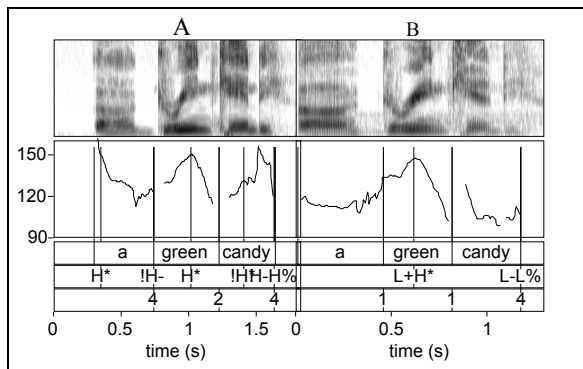


Figure 2: *Example F0 contours for adjective-noun pairs: A) Contrastive Noun and B) Contrastive Adjective.*

A small number of utterances (10) provided the opportunity to examine contrastiveness defined at the D-level but not at the DS-level. For these trials, a discourse segment boundary intervened between the two mentions of the target word (e.g. utterance 069.5 *blue house* and 095.3 *blue bell* in Table 1). Of two adjectives in contrastive contexts at the D-level, one bore a L+H\* pitch accent. Of eight nouns in contrastive contexts at the D-level, one bore a L+H\* accent, one was deaccented, and the remaining six bore another type

of accent. Although this pattern is not conclusive, it suggests that contrastiveness defined at the D-level may lead to an accentuation on the contrastive word across DS-boundaries.

#### 4. Conclusions

The present study provides the preliminary description of prosodic properties of unscripted speech that was collected though a carefully designed cooperative task. The finer-grained discourse analysis demonstrates that deaccentuation is better predicted by the *given* status of the word defined at the local discourse segmental (DS) level than by *givenness* defined at the global discourse (D) level. The data also suggest that word position has a large effect on accent distribution, e.g., deaccentuation related to *givenness* is more common for the noun than for the adjective position. The occurrence of certain accent type such as L+H\*, was indeed predictable from certain informational status such as *contrastiveness*, although more conclusive generalizations about the distribution of certain accent types should wait the result of the analysis of the other subjects' utterances. Since the experimental technique introduced in the present study may allow the analysis of listeners' reaction to certain accent types in natural discourse context, future research shall investigate the effect of accent distribution in speech comprehension.

#### 5. References

- [1] Grosz, B. J.; Sidner, C. L., 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12 (3), 175-204.
- [2] Chafe, W., 1976. Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In C. Li (Ed.), *Subject and topic* (pp. 25-56). New York: Academic Press.
- [3] Cruttenden, A., 1986. *Intonation*. Avon: Cambridge University Press.
- [4] Bard, E.G.; Aylett, M.P., 1999. The dissociation of de-accenting, givenness, and syntactic role in spontaneous speech. In *Proceedings of the Fourteenth International Conference of Phonetic Sciences, San Francisco*, 1753-1756.
- [5] Terken, J.; Hirschberg, J., 1994. Deaccentuation of words representing "given" information: Effects of persistence of grammatical function and surface position. *Language and Speech*, 37, 125-145.
- [6] Nakatani, C. H., 1997. *The computational processing of intonational prominence: A functional prosody perspective*. Ph.D. dissertation. Harvard University.
- [7] Pierrehumbert, J. B., 1980. *The phonology and phonetics of English intonation*. PhD dissertation. Massachusetts Institute of Technology.
- [8] Pierrehumbert, J.; Hirschberg, J., 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication*, P. Cohen, J. Morgan, M. Pollack (eds.). Cambridge: MIT Press, 342-365.
- [9] Beckman, M. E.; Ayers, G. M., 1994. *Guidelines for ToBI labelling, vers 2.0*. Ohio State University.
- [10] Boersma, P. W.; Weenick, D., 1992-2003. Praat doing phonetics by computer (Version 4.1.12).