

The probabilistic relationship between pitch accents and information status in public speech

Suyeon Im¹, Jennifer Cole², Stefan Baumann³

¹University of Illinois at Urbana-Champaign

²Northwestern University ³University of Cologne

sim16@illinois.edu, jennifer.cole1@northwestern.edu, stefan.baumann@uni-koeln.de

Abstract

Pitch accents encode semantic or pragmatic meaning in English [1], [2]. This study examines the relationship between pitch accent assignment and information status (IS), adopting the richer IS scheme of RefLex [3], in an intact sample of public speech from a TEDTalk. 361 words from the speech sample were annotated for IS specified in terms of referential, lexical, and alternative (focus) conditions. Results show different effects of referential vs. lexical givenness on accent assignment. Only referential givenness has the expected effect of given words being (mostly) unaccented. The TEDTalk speaker uses accent differently from what has been reported in prior work [4], with a much more variable distribution of accent across IS conditions, and an overall weaker probabilistic association between accent and IS. This study demonstrates the necessity of distinguishing lexical and referential givenness, and the effect of speech style on prosodic variability.

Index Terms: prosodic prominence, pitch accent, information status, speech style

1. Introduction

In English, pitch accents may convey the meaning of a word in relation to its discourse context [1], [2]. Pitch accents are generally assigned to words that carry new information and to focused words whose referent is identified in relation to a restricted set of alternatives. Words that are previously mentioned and whose referents are in the common ground often remain unaccented. Within the prosodic phrase, the tonal melody of the nuclear pitch accent (in rightmost position) has been found to distinguish different IS conditions: H* is used for new information and L+H* for scaled interpretation in discourse [4].

There are several known and potential sources of variation in pitch accent assignment. For example, speaker variability in the production of intonational features is reported [5], [6], but the extent of variation in accent assignment across a range of IS conditions has not been investigated. Variation in the use of accents by speech style has likewise not been widely investigated. Much of the research on the relationship between pitch accent and IS has focused on nuclear pitch accents, while prenuclear accents are little studied, but claimed to be “ornamental” [7] or rhythmic [4]. IS conditions on accent assignment have focused mainly on distinctions in lexical givenness (i.e., prior occurrence of lexical items), without considering the possibility of differential influence of referential givenness. A final observation is that accent assignment has been most often investigated with experimental methods using minimally contextualized

utterances (e.g. question-response prompts, picture description task), or in excerpts from conversational speech, which do not fully capture the richer discourse context that occurs in intact samples of extended, natural discourse.

In this paper we examine the association of pitch accents with IS, assuming a more elaborated model of IS, and within a complete, coherent discourse context. We adopt the RefLex scheme introduced for the analysis of German [3], which differentiates IS conditions related to referential, lexical, and focus (alternative reference) conditions. We present an exploratory study of prosody in a public, motivational speech style, through the analysis of a speech from a TEDTalk. We wish to answer the following three questions about the relationship between accent and IS in this speech sample: (1) Does accent assignment distinguish words that potentially carry IS from words that do not? (2) Does accent assignment distinguish words that are given from those that are not? (3) Does accent type distinguish among different categories of IS?

2. Methodology

The speech material is obtained from a TEDTalk titled “Try Something New for Thirty Days” [8], which is a motivational speech delivered by a male speaker of American English, in a clear speech style ($t = 2'25''$). The sample contains 361 words. Pitch accents were annotated by the second author and a second trained labeler following ToBI annotation conventions [9]. IS was annotated by the first and third author using the RefLex scheme [3], [10]. RefLex labels for referential conditions (r-) are assigned to individual nouns or entire noun phrases, and labels for lexical (l-) and alternative (alt-) conditions are assigned to all content words. Table 1 shows the basic RefLex scheme used in the study, which is a simplified version of the full RefLex scheme [10]. The labels are described in the decreasing order of “givenness” in each condition: e.g. r-given > r-bridging > r-unused > r-new.

Table 1: *Basic RefLex scheme*

Level	Label	Description
r-level	r-given	Referents present in previous discourse context
	r-bridging	Discourse-new entities that depend on context
	r-unused	Globally unique entities that are discourse-new
	r-new	Non-unique discourse-new entities
l-level	l-given	Active (salient) concepts
	l-new	Inactive concepts
alt-level	alt	Alternative concepts

3. Predictions

Accent assignment is predicted to distinguish words that *potentially* carry discourse information from words that do not. Examples of the former are content words, while the latter include function words. The words that carry discourse information are potentially accented while the words that do not are ineligible for accent. This follows from the standard assumption that pitch accents encode the meaning of a word in discourse context [1], [2] and should be the case for all three IS conditions examined here (*r*-, *l*-, *alt*).

Accent assignment is also expected to distinguish words that are given from those that are not: the words with given information should be unaccented while the words with accessible or new information should be accented. This follows from the assumption that pitch accents mark levels of new information [1]. Words that are eligible to carry discourse information but which are referentially or lexically given in a particular discourse context should be unaccented.

The tonal type of pitch accent is predicted to distinguish among different categories of IS [4]. H^* is the pragmatically neutral accent and as such is expected to be found with any IS categories except given [1]. $L+H^*$ is described as signaling a referent that is one among alternatives on a salient scale or as inferred from the discourse context, and is typically used to mark contrastive or corrective focus [2]. $L+H^*$ also has a higher degree of perceived prominence than H^* [11]. Accordingly, we expect $L+H^*$ to be associated with the IS categories *r*-new and *alt*. $!H^*$ is used to enumerate objects in succession (e.g. shopping list) [12], and is expected to be used to mark accessible information, with *r*-bridging and *r*-unused. L^* is expected to be found with given categories since it is to be used with items that are salient but are not to be added to a predication made by the speaker [2].

4. Results

We present results addressing the three research questions introduced in section 1, about the accent status (accented or unaccented) of words that are eligible vs. ineligible to encode IS; the accent status of words that are referentially or lexically given; and the type of pitch accents that occur under different conditions of IS.

The distribution of pitch accents in each IS condition are shown in Table 2. For some additional accent types, less than ten tokens were obtained ($H+!H^* = 1$, $L^*+H = 3$) and these items were re-assigned to other accent types with a similar contour shape (H^* for $H+!H^*$, $L+H^*$ for L^*+H). The RefLex scheme includes some additional IS categories for which few tokens were obtained (*r*-cataphor = 1, *l*-accessible = 7), so these IS categories were merged with others (*r*-bridging for *r*-cataphor, *l*-given for *l*-accessible).

As can be seen from Table 2, in all IS levels the greatest number of unaccented words occur in the “none” categories, i.e., with words that are not assigned to any IS categories in the RefLex scheme, and that are considered as ineligible to convey meaning related to the specific discourse context. We note that the “none” categories do include some accented words, but they are infrequent. Among those words labeled as given, there is a difference between referential and lexical conditions. Words in the *r*-given condition follow the expected pattern of being mostly unaccented, though some *r*-given words are accented. In the *l*-given condition more words are accented than unaccented.

Table 2: Distribution of pitch accents by IS categories

Level	Label	Unacc.	L^*	$!H^*$	H^*	$L+H^*$
<i>r</i> -level	<i>r</i> -none	124	5	15	31	26
	<i>r</i> -given	34	2	2	11	12
	<i>r</i> -bridging	5	1	6	5	4
	<i>r</i> -unused	3	5	6	11	8
	<i>r</i> -new	11	6	4	9	15
<i>l</i> -level	<i>l</i> -none	152	1	7	27	20
	<i>l</i> -given	9	5	8	11	5
	<i>l</i> -new	16	13	18	29	40
<i>alt</i> -level	<i>alt</i> -none	164	17	31	61	53
	<i>alt</i>	13	2	2	6	12

Pearson’s chi-squared test with Yates’ continuity correction is used to test whether the assignment of pitch accents distinguish words that potentially carry IS from words that do not, and to distinguish words that are referentially or lexically given from those that are not given. For the first comparison, the analysis is based on the counts of unaccented and accented (L^* , $!H^*$, H^* , $L+H^*$) words in relation to the “none” and “all” categories of IS. The “all” category combines all labeled IS conditions except “none”: *r*-all includes *r*-given, *r*-bridging, *r*-unused, and *r*-new; *l*-all includes *l*-given and *l*-new; *alt*-all includes only *alt*. For the second comparison the analysis is based on the counts of unaccented and accented (L^* , $!H^*$, H^* , $L+H^*$) words as a function of the words with “given” and “non-given” categories of IS. The “non-given” category includes all IS categories except “given” and “none” categories in each condition (*r*-none-given for *r*-bridging, *r*-unused, *r*-new; *l*-none-given for *l*-new). The alternative condition is not analyzed for the analysis of “given” words, since it does not have a given category. The results are shown in Table 3 below. The assignment of pitch accent significantly distinguishes words that carry IS from words that do not, in both referential and lexical conditions. Among IS categories, the assignment of pitch accents significantly distinguishes words that are given from those that are not in the referential condition only.

Table 3: Chi-squared values for categories as carrying IS or not (“all/none”), and givenness or not (“given/non-given”)

Level	All/None			Given/Non-given		
	χ^2	df	p	χ^2	df	p
<i>r</i> -level	27.96	1	<.01	21.14	1	<.01
<i>l</i> -level	113.32	1	<.01	1.40	1	.24
<i>alt</i> -level	1.70	1	.19	NA	NA	NA

As a main finding, we conclude that pitch accents are primarily assigned to words that are eligible to carry referential and lexical givenness in discourse. Within the categories of IS, there is a difference between referential and lexical givenness. Pitch accents are assigned to the words that convey referentially accessible and new information but are much less likely to occur with words that carry referentially given information. In comparison, the assignment of pitch accents does not distinguish words that are lexically given from those lexically new. These findings support the RefLex model in distinguishing referential and lexical givenness. We find that the assignment of pitch accents is probabilistic, since some words with “none” and “given” categories do receive

pitch accents, and all types of pitch accents occur with these IS categories.

We are also interested in whether individual types of pitch accent encode different IS categories. p -values were obtained from Fisher's Exact Test based on 2000 replicates since some cells have less than 5 observations. The test is based on the counts of each type of pitch accent (L*, !H*, H*, L+H*) as a function of different IS categories, excluding the r-none and l-none categories. (Alt-level is not examined since it has only one category--*alt*). The results show that there is no significant relationship between accent type and IS category (for r-level: $p = .29$, l-level: $p = .10$).

Figures 1-5 show the percent and the frequency of different types of pitch accent as a function of IS categories in each condition (Figure 1-2 for r-level; Figure 3-4 for l-level; Figure 5 for alt-level).

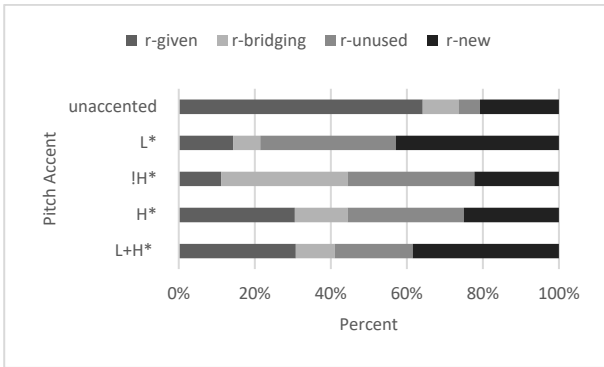


Figure 1. Percent of pitch accent types as a function of referential givenness in the public speech sample

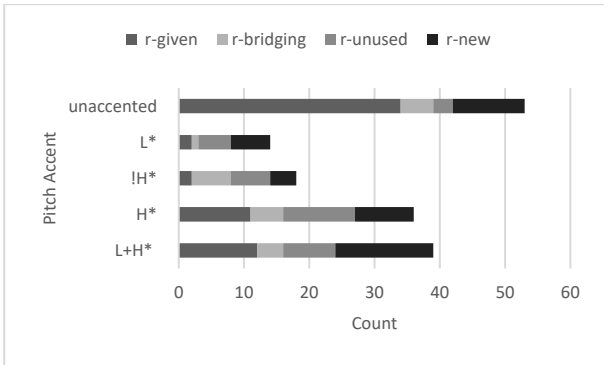


Figure 2. Frequency of pitch accent types as a function of referential givenness

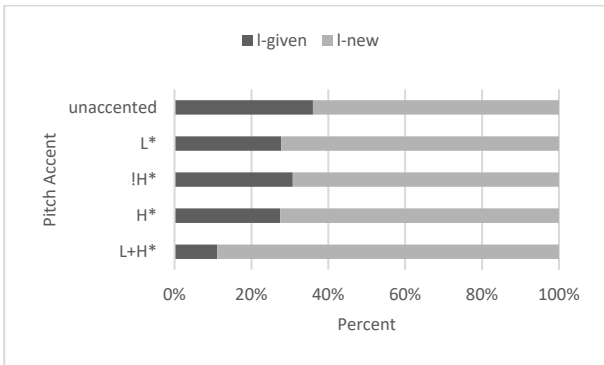


Figure 3. Percent of pitch accent types as a function of lexical givenness

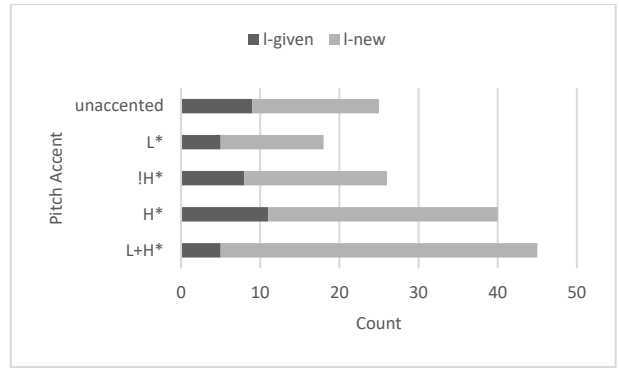


Figure 4. Frequency of pitch accent types as a function of lexical givenness

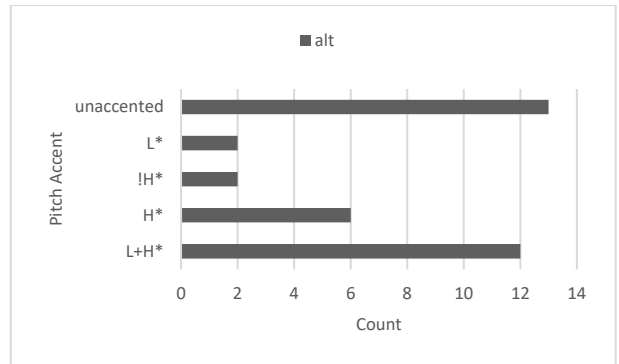


Figure 5. Frequency of pitch accent types as a function of alternative concepts

From Figures 1-2, there is a numeric trend that !H* is the most frequent pitch accent for r-bridging, and L+H* for r-new. We set L* aside for r-new since when L* is a nuclear pitch accent, it is found to be frequently used for indicating continuation in conjunction of the following high boundary tone by the speaker in this public speech sample. R-unused is encoded mostly by H* although it is also frequently marked by other pitch accents. From Figures 3-4, it is observed that L+H* is the most frequent pitch accent for l-new. Finally, Figure 5 shows that L+H* is the most frequent pitch accent type for marking the presence of alternative concepts.

We find that all pitch accent types are used in all conditions, and any association between pitch accents and IS is probabilistic. This is support for the findings reported e.g. for Neapolitan Italian [13] and German [3], [14] from production data and for German from perceptual goodness ratings [15]. The speaker also strongly favors L+H* for the alternative condition, as predicted. Despite of this, he occasionally uses other accent types as well. In this respect, our model speaker uses accent differently from what has been reported in prior work [4], where different pitch accents encode different IS conditions.

5. Discussion and conclusion

This study shows that the relationship between the assignment of pitch accents and IS conditions is probabilistic, in alignment with previous findings, e.g. for Neapolitan Italian and German. There is little evidence in our public speech that the type of pitch accent predicts the meaning of a word in relation to the discourse context, refuting the notion of a strict one-to-one mapping between accent type and IS condition.

Among IS conditions, there is a difference between referential and lexical conditions in that only referentially given words are probabilistically unaccented. This supports the RefLex model distinction of two levels of givenness. In this speech sample, which is representative of a motivational and public speech style, accenting patterns are very different from what has been reported for conversational speech. All accent types are used in each referential and lexical condition. Although contrastive focus probabilistically favors L+H* over other accents, this type is also widely used in other IS conditions, and is surprisingly frequent in r-given.

To further explore the prosodic style of the speech sample analyzed here, we compare this sample (from one male speaker) to a sample of conversational speech from eight male speakers from the Buckeye corpus [16]. Comparisons are made for accent assignment and max f0 from the accented word, and for prominence ratings by American English speakers as reported in [17]. From Figure 6, we observe that the public speech has a higher occurrence of accented words with over half of the words accented, more frequent use of L+H*, and higher max f0 than in the Buckeye speech data. It is interesting to note that despite the frequent accenting, the prominence rating by listeners is similar between the public and the conversational speech. This shows that accenting patterns can be varied across different speech styles although they may not affect the judgment of prominence perception by untrained listeners.

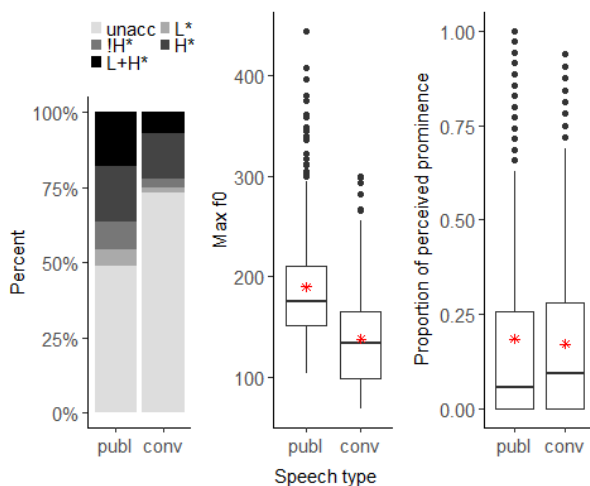


Figure 6. Comparison between public speech (publ) and conversational speech (conv) from Buckeye corpus.

This study is to our knowledge the first one to demonstrate the relationship between accent assignment and IS conditions in speech produced in a public speech style. Our findings provide evidence that models of the association between accenting and IS should distinguish between lexical and referential givenness, and should consider speech style. The stylistic use of accenting respects the probabilistic pattern of reserving accent for words that are eligible to bear discourse information. It also respects the pattern of not assigning accent to given words, but only for referential givenness. The most significant feature of this speaker's prosodic style is the complete disassociation of accent type and referential IS condition. He varies accent type freely, with little regard for IS condition, and strongly favors the more prominent accent types – especially L+H*. In future and on-going work, we intend to examine the perception of prosodic prominence in relation with expectation-driven (discourse meaning) vs.

signal-driven factors (pitch accents and acoustic cues) in this style of public speech.

6. Acknowledgements

The research reported here was supported by NSF BCS 12-51343.

7. References

- [1] D. R. Ladd, *Intonational Phonology*. Cambridge University Press, 2008.
- [2] J. Pierrehumbert and J. B. Hirschberg, "The meaning of intonational contours in the interpretation of discourse," *Intentions in Communication*, pp. 271-311, 1990.
- [3] S. Baumann and A. Riester, "Coreference, lexical givenness and prosody in German," *Lingua*, vol. 136, pp. 16-37, 2013.
- [4] S. Calhoun, *Information Structure and the Prosodic Structure of English: A Probabilistic Relationship*. Doctoral dissertation, University of Edinburgh, 2007.
- [5] T.-J. Yoon, "Speaker consistency in the realization of prosodic prominence in the Boston University Radio Speech Corpus," in *Speech Prosody 2010 – Fifth International Conference*, 2010.
- [6] G. Kochanski, E. Grabe, and J. Coleman, "The difference between a question and a statement: A cross-dialect survey," *The Journal of the Acoustical Society of America*, vol. 115, no. 5, 2004.
- [7] D. Büring, "Intonation, semantics, and information structure," in Ramchand, G.; Reiss, C., editors. *The Oxford handbook of linguistic interfaces*. Oxford: Oxford University Press, 2007, pp. 445-473.
- [8] M. Cutt, Try Something New for Thirty Days, TED. <https://ed.ted.com/lessons/try-something-new-for-30-days-matt-cutt>
- [9] M. E. Beckman and J. Hirschberg, *The ToBI Annotation Conventions*. Ohio State University, 1994.
- [10] A. Riester and S. Baumann, *The RefLex Scheme – Annotation Guidelines*. Vol. 14 of *SinSpeC*, Working Papers of the SFB 732. University of Stuttgart, 2017.
- [11] J. I. Hualde, J. Cole, C. Smith, C. Eager, T. Mahrt, and R. Napoleão de Souza, "The perception of phrasal prominence in English, Spanish and French conversational speech," *Proceedings of the International Conference on Speech Prosody*, vol. 8, pp. 459-463, 2016.
- [12] Transcribing Prosodic Structure of Spoken Utterances with ToBI, MIT. <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/exercises/>
- [13] F. Cangemi and M. Grice, "The importance of a distributional approach to categoriality in autosegmental-metrical accounts of intonation," *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, vol. 7, no. 1, 2016.
- [14] D. Mücke and M. Grice, "The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation?," *Journal of Phonetics*, vol. 44, pp. 47-61, 2014.
- [15] F. Cangemi, M. Krüger, and M. Grice, "Listener-specific perception of speaker-specific production in intonation," in Fuchs, S.; Pape, D.; Petrone, C.; Perrier, P., editors. *Individual Differences in Speech Production and Perception*. Frankfurt am Main: Peter Lang, 2015, pp. 123-145.
- [16] M. A. Pitt, K. Johnson, E. Hume, S. Kiesling, and W. Raymond, "The Buckeye corpus of conversational speech: Labeling conventions and a test of transcriber reliability," *Speech Communication*, vol. 45, no. 1, pp. 89-95, 2005.
- [17] J. Cole, T. Mahrt, and J. I. Hualde, "Listening for sound, listening for meaning: Task effects on prosodic transcription," *Proceedings of the International Conference on Speech Prosody*, pp. 859-863, 2014.