

Ambiguity in prominence perception in spoken utterances of American English

L.C. Dilley and S. Shattuck-Hufnagel

Speech Communications Group, RLE, 36-511, MIT, 77 Massachusetts Avenue, Cambridge, Mass. 02139

Abstract: Ambiguities in whether a syllable in a spoken utterance is prominent may arise in contexts where full-vowel syllables occupy adjacent locations. We explored whether contextual factors could influence prominence perception in the phrase *they're all right now*, which contains a morphosyntactic ambiguity. This phrase was uttered with either *maybe* (which is lexically strong-weak, S-W) or *for sure* (W-S) preceding. Comparison stimuli were created by cross-splicing each initial context (*maybe* vs. *for sure*) to the *they're all right now* of the other utterance. Thirteen subjects judged prominence for syllables in 96 sentences (four repetitions of each of eight experimental sentences and 16 control sentences). Of the nine subjects who indicated sensitivity to lexical stress, the number of judgments of prominence on *they're all right now* differed significantly in five out of eight conditions depending on whether the preceding context was *maybe* or *for sure*. Results support the hypothesis that context can influence prominence judgments, and that portions of spoken utterances require interpretation.

INTRODUCTION

This paper explores whether perception of prominence may sometimes be *ambiguous*. Anecdotally, we have noticed that listeners are prone to uncertainty in deciding which syllable or syllables are prominent in tasks involving prosodic transcription of speech. One such context occurs when two full vowel syllables occupy adjacent positions and the lexical stress information is not specified in the lexicon. Ambiguously stressed words such as *digest* or proper names like *Mullins* contain adjacent full vowel syllables; such words may pose difficulties for the phonological system of the language (Shattuck-Hufnagel 1995; Huss 1978).

Adjacent full vowel contexts may also occur *across words*. The phrase *they're all right now* contains a morphosyntactic ambiguity. We hypothesized that if a context of unambiguous lexical stress (W-S or S-W) were to precede *they're all right now*, it would affect judgments of which syllables were prominent. Specifically, we suspected that if a S-W context (such as the word *maybe*) preceded the phrase, that the listener would prefer to hear a S-W-S-W on *they're all right now*, resulting in greater likelihood that *they're* and *right* would be heard as prominent. Moreover, we hypothesized that in the case of a preceding "weak-strong" context (as in *for sure*) that the listener would prefer to hear W-S-W-S, resulting in greater likelihood that *all* and *now* would be judged prominent.

METHOD

The sentences *maybe they're all right now* and *for sure they're all right now* were uttered using the intonation shown in Figure 1.¹ The intonation contour was chosen because it was hypothesized that its repeated nature (alternating peaks and valleys on each syllable) would reinforce the expectation of *either* a S-W-S-W or a W-S-W-S pattern, depending on the preceding context. These sentences were produced (i.e. intended by the speaker) to have the same F_0 contour, but with prominences on *may-*, *they're* and *right* (i.e. the peaks) in *maybe they're all right now* and on *sure*, *all*, and *now* (i.e. the valleys) in *for sure they're all right now*. The original two utterances were then cross-spliced to create two experimental utterances with the initial context *maybe* concatenated with the version of *they're all right now* originally produced as part of the utterance beginning *forsure...*, and vice versa. This is shown in Figure 1.

Sixteen control sentences, e.g. *perhaps they're almost done*, which contained no syntactic ambiguity, were produced with the same intonation contour and digitized. An audio tape was prepared which consisted of four repetitions of each of the 24 sentences (8 experimental plus 16 control), for a total of 96 sentences.

¹Utterances used as stimuli in this experiment are available via anonymous ftp to lexic.mit.edu.

Thirteen subjects between the ages of 20-45 with normal hearing who were native or near-native speakers of English were asked to underline the syllables which sounded prominent or emphasized to them.

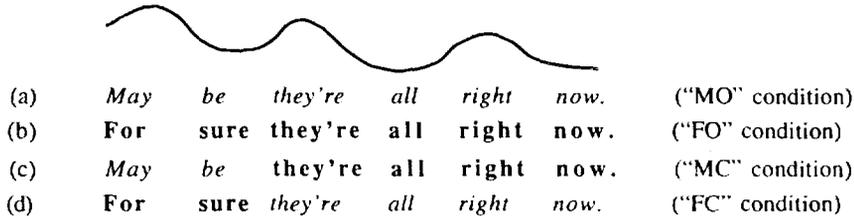


FIGURE 1. Experimental stimulus conditions to test the effects of preceding context on judgments of prominence. (a) and (b) show the original sentences (conditions "MO" and "FO") while (c) and (d) illustrate concatenated conditions ("MC" and "FC").

RESULTS

Four out of thirteen subjects frequently indicated lexically unstressed syllables (e.g. *per-* in *perhaps*) as prominent. Because we were interested in lexical stress effects, we used data from the nine subjects who indicated lexical stress correctly most of the time. Because the number of data points per subject was small, we compared prominence judgments for *they're all right now* across subjects. Each subject had eight opportunities to judge a word's prominence in a particular stimulus condition ("MO"/"FO"/"MC"/"FC"), for a total of $n = 72$ prominence judgments per word per condition. To review our hypotheses, we suggested that *they're* and *right* would be more likely to be heard as prominent when preceded by *maybe*, and that *all* and *now* would be more likely to be heard as prominent when preceded by *for sure*. Accordingly, we compared judgments the "MO" condition with judgments in the "FC" condition, and judgments in the "FO" condition with those in the "MC" condition, respectively. A one-sided test of significance for these hypotheses was performed and results are given in Table 1. Differences in prominence judgments in the "FO/MC" case which did not reach significance were nevertheless in the direction suggested by the hypotheses.

Stimulus conditions	Prominence condition			
	<i>they're</i>	<i>all</i>	<i>right</i>	<i>now</i>
MO/FC	YES $p < 0.001$	YES $p < 0.025$	YES $p < 0.001$	YES $p < 0.025$
FO/MC	YES $p < 0.001$	NO	NO	NO

TABLE 1. Significance levels for differences in number of judgments of "prominent" for *they're*, *all*, *right*, and *now* in different stimulus conditions.

DISCUSSION AND CONCLUSIONS

The results support the hypothesis that prominence perception may be influenced by contextual factors. The variability in prominence judgments by subjects is consistent with the idea that some portions of speech may be *ambiguous* with respect to prominence. The contribution of various factors (lexical stress, intonation, etc.) to resolving ambiguity in prominence perception warrants further inquiry. A more detailed discussion of ambiguity in prominence perception and of hypotheses presented here are given in Dillely (in preparation).

ACKNOWLEDGEMENTS

This work was supported by NIH grants DC 0075-32 and DC 02125-01, and NSF subcontract C-Q0198 with SRI International, as well as funding from the MIT Undergraduate Research Opportunities Program.

REFERENCES

- Dillely, L. (1998) "A perceptual study of rhythmic and intonational regularity," *MIT Speech Group Working Papers*, in preparation.
- Huss, Volker (1978) "English word stress in the post-nuclear position," *Phonetica* 35, pp. 86-105.
- Shatzk, Ulf-Fredrik, S. (1995) "Pitch accent patterns in adjacent stress and alternating stress words in American