ON THE PROMINENCE OF ACCENT IN STRESS REVERSAL

Jennifer Cole, José I. Hualde, Christopher Eager, Timothy Mahrt

University of Illinois
{jscole | jihualde | eager2 | tmahrt2}@illinois.edu

ABSTRACT

We investigate the prominence of English words with stress reversal (e.g. *èlevátion* 2-1 → *élevàtion* 1-2). We ask what motivates the occurrence of the "early high" (1-2) pattern outside of stress clash contexts, and consider the hypothesis that it marks prominence non-locally. Experiment 1 tests the effect of prominence pattern on memory. Given its markedness and location at phrasal onset, we hypothesize that early high pitch broadly facilitates recall for sentence information. This hypothesis is not confirmed, suggesting that the effect of pitch accent on memory may be restricted to the accented word. In Experiment 2 listeners perform a prominence-rating task on the same patterns. Results show that early high is prominence-lending, but with weaker prominence than the lexical (2-1) stress pattern. The combined findings suggest a hybrid function for early high in marking the beginning of a discourse-level prosodic unit, and in lending prominence to the early high-accented word.

Keywords: stress reversal, prosody, English stress, prominence perception.

1. INTRODUCTION

In English, words where the syllable with primary stress is preceded by another syllable with secondary stress (2-1 words, e.g. *èlevátion, òptimístic*) allow a less frequent pronunciation with reversal of prominence (e.g. *èlevátion* → *élevàtion*), in which case their pattern becomes identical to that of 1-2 words (e.g. *élevàtor, súpermàrket*). Stress reversal is produced by the association of a high-tone pitch-accent with the syllable that bears secondary stress in the unmarked pronunciation of the word. We will thus refer to this phenomenon as "early high" (see [1-3]). Early high in 2-1 words has been reported in stress clash contexts ([1-7]), but is not limited to this context. We notice that this pattern is frequent in "news broadcaster style", most notably at the start of a topic unit, and it also appears in conversational contexts, in both cases not restricted to contexts of stress clash. A question that arises is what motivates the application of early high outside of contexts of stress clash. Here we consider the possibility that early high functions like other English H-tone pitch accents in marking a constituent as new information, but with increased salience due to its non-canonical location on the syllable with secondary stress. The early high pattern places a H-tone pitch accent on the first stressed syllable in the prosodic phrase, thus also serving a demarcative function, signaling the beginning of the constituent that introduces new information. In other words, the early high pattern may tell the listener to "*listen up—important information follows!*". Lending support for a demarcative function of early high is the observation that prominence reversals seem not to occur when the secondary stress follows the primary stress in the word. We do not find words like *élevàtor* (1-2) realized as *èlevátor* (2-1), regardless of information status or position in the prosodic phrase. Thus, an important property of stress-reversal marked with early high is the location of the pitch accent at the start of a prosodic phrase, frequently coinciding with the start of a discourse topic.

This paper reports perceptual evidence from two experiments testing the status of early high as prominence-lending and demarcative. Experiment 1 tests the effect of pitch accent patterns on memory. Other work has suggested that pitch-accent type affects recollection: words having a salient pitch accent (L+H*) are remembered better [8]. We hypothesize that a word with a lexical 2-1 pattern has greater perceptual salience when realized with the early high accent pattern (1-2) than with the canonical pitch accent pattern (2-1), or when unaccented. And in its demarcative function at the sentence beginning, early high may call listeners' attention not only to the early high accented word, but also to the information to follow in the prosodic phrase [see 9]. Accordingly, we predict that listeners will be more accurate in recalling sentence information following an early high accented word.

Experiment 2 uses a prominence labeling task to assess the perceived prominence of early high in words like *elevation* against canonical (1-2) and unaccented prominence patterns.

2. MATERIALS AND PARTICIPANTS

2.1 Materials

Speech stimuli used in both experiments were drawn from the same set of materials, described here. 30

words with a lexical 2-1 stress pattern were chosen to form sets of three sentences each, for a total of 90 experimental sentences. In each sentence set the critical word with the 2-1 stress pattern was the first content word of the complex subject NP, and three different sentences were formed from by varying the words in the continuation of the subject NP and the following VP. For example, the sentence set in (1) is based on the critical 2-1 word *university*. The sentence sets were constructed to produce triplets of sentences that begin with the same word sequences, which should make the recall of any one sentence in Exp. 1 (section 3) potentially more difficult.

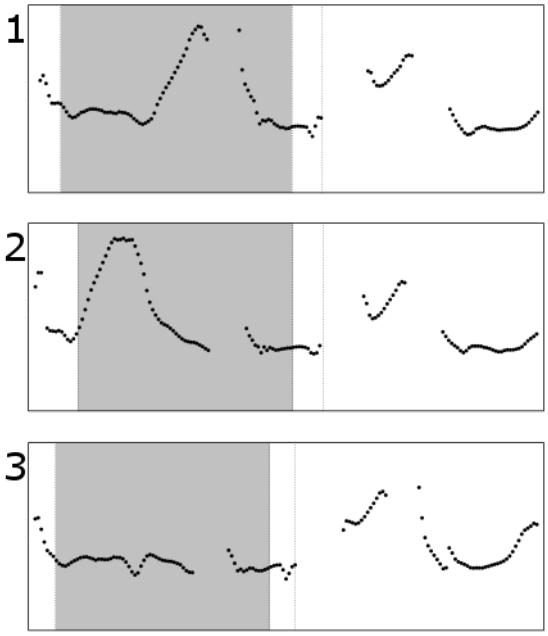
1a. [*The University of Kosovo*]$_{NP}$ *won the tournament that nobody thought they could win.*
1b. [*The University of Katmandu*] $_{NP}$ *had a budget crisis due to embezzlement and mismanaged funds.*
1c. [*The University of Kenya*] $_{NP}$ *built a new library on top of archeological ruins.*

Sentences in these sets were further manipulated to introduce three prosodic patterns over the initial content word: Accent Pattern (hereafter AP) 1 = H* accent on syllable with lexical primary stress, AP2 = early high, and AP3 = unaccented. The three prosodic patterns for subject NPs are illustrated in Figure 1. These three distinct patterns were produced for each sentence, for a total of 270 distinct utterances (30 sets x 3 sentences x 3 prosodic patterns).

A female speaker of American English (the author JC) was recorded to create the 270 stimulus utterances from these sentence materials, as follows. The subject NPs from the three sentences within a set were recorded separately from their VP continuations, for greater uniformity in the naturally produced prosodic patterns over the 90 subject NPs. Each subject NP (e.g., *The University of Kosovo*) was produced in isolation as a full prosodic phrase, in each of the three different prosodic patterns. The three differently accented versions of each subject NP were then spliced onto the same VP continuation (always from AP1), resulting in three prosodically distinct versions of each sentence, as illustrated in (2) for the sentence in (1a).

2a. [*The Ùnivérsity* (2-1) *of Kosovo*]$_{NP}$ *won the tournament that nobody thought they could win.*
2b. [*The Únivèrsity* (1-2) *of Kosovo*]$_{NP}$ *won the tournament that nobody thought they could win.*
2c. [*The University* (unaccented) *of Kosovo*]$_{NP}$ *won the tournament that nobody thought they could win.*

Figure 1. Pitch tracks from three different productions of the NP *The University of Kosovo* (sentence 1a, above) with the word *University* highlighted in grey. In AP1 (top panel) there is a High-tone pitch accent on the primary stressed syllable; in AP2 (middle panel) there is a pitch accent on the secondary stressed syllable (ie., early high); in AP3 (bottom panel) there is no accent (unaccented).



The 270 spliced utterances were divided into 3 lists such that each list contained one of the APs for each of the 90 lexically distinct sentences, and a different AP for each of the three sentences from the same set. The assignment of utterances to the three lists is illustrated in Table 1 for the nine prosodically distinct utterances from the *University* sentence set in (1).

Table 1. The assignment of sentence stimuli to three lists, illustrating the distribution of sentence continuations and APs across lists. The 9 lexically and prosodically distinct sentences from set 1 are represented by their subject NPs and AP (1-2, 2-1, and unaccented (UA)) of the critical word, *University*. The shaded cells represent the utterances that were chosen for the prominence labeling task and were also used as the test set for the recall task.

List 1	List 2	List 3
2-1: Ùnivérsity of Kosovo	*2-1:* Ùnivérsity of Katmandu	*2-1:* Ùnivérsity of Kenya
*1-2:*Únivèrsity of Kenya	*1-2:*Únivèrsity of Kosovo	*1-2:*Únivèrsity of Katmandu
UA: University of Katmandu	*UA:* University of Kenya	*UA:* University of Kosovo

A set of 30 utterances were chosen from each list as the test set, including one sentence for each of the 30 critical words (shaded cells in Table 1). The recall task (Exp. 1) used all 90 utterances from each list,

with the 30 sentences from the test set chosen as the recall test items (see section 3.1). The prominence labeling task (Exp. 2) used only the 30 utterances from the test set of Exp. 1.

2.2. Participants

30 college students participated in this study (21 females, 9 males). All participants self-identified as native, monolingual speakers of American English with no impairments in speaking, hearing or reading. Participants had no prior background in phonetics or phonology and were not trained in prosodic transcription. Participants were randomly assigned to one of the three stimuli lists (Table 1), with 10 participants assigned for each list. Every participant heard each of the 90 experimental sentences only once, heard the three sentences from a sentence set with a different AP on each one, and heard an equal number of utterances with APs 1, 2 and 3.

3. RECALL EXPERIMENT (EXP. 1)

3.1. Methods

The 90 utterances in each list were presented to participants through headphones in 6 blocks of 15 utterances. Within each block, utterances were presented in random order, with successive utterances separated by a short tone. The 3 sentences from the same set (e.g., 1a-c), each with a different AP, were presented in the same block, separated by one or more unrelated sentences. For instance, the first block in one of the lists contained 3 sentences with each of the following 5 beginnings: *The information for…, The revolution in…, The University of…, Opportunities for…* and *Our understanding of...* The three APs were distributed evenly within and across blocks.

After each block, participants were asked to answer 5 multiple-choice questions that tested recall on 5 sentences from the test set. These questions were given in written form and consisted of sentences that they had heard in the preceding block but with a blank following the first content word and with 3 response options. For example, following the block containing the three sentences in set (1a-c), subjects were presented with the test item in (3), for which the correct response would be *Kosovo* (see 1a).

(3) **Test item:**
The University of _____ won the tournament that nobody thought they could win.
Response options: *Kosovo; Katmandu; Kenya*

3.2. Results

The effect of AP on recall accuracy was tested with mixed effects logistic regression in R [10] using glmer() in the lme4 package [11] with accuracy as the response and AP as a fixed effect coded with sum contrasts. Following [12], the maximal random effects structure supported by our data is random intercepts for question and transcriber:

Table 2. Mixed Effects Logistic Regression on Recall Accuracy (Residual Deviance: 1152.7 on 895 df)

Fixed Effects	**Estimate**	**SE**
Intercept (Grand Mean)	0.4878	0.1691
AP1 (2-1)	0.1113	0.1034
AP2 (Early High, 1-2)	-0.2364	0.1026
Random Effects	**SD**	
Question (intercept)	0.5089	
Transcriber (intercept)	0.6579	

The effect of accent pattern was not significant (parametric bootstrap test using mixed() in the afex package [13] with 10,000 simulations: $p = .069$), and the output in Table 2 shows that Early High trends toward lower recall accuracy than the grand mean. In other words, it is not the case that subjects in this experiment were more accurate remembering information from sentences where the first content word had an early high pattern.

4. PROMINENCE LABELING (EXP. 2)

4.1. Methods

The same 30 participants took part in Experiment 2, a prominence labeling task using the method of Rapid Prosody Transcription (see [14]), which followed Experiment 1. Participants listened to the 30 utterances from the test set of their assigned stimulus list (the same utterances used as test items in Exp. 1), and were instructed to identify the words they perceived as prominent. Prominence was judged based only on auditory impression. The exact instructions participants received were the following: *"In normal speech, speakers pronounce some word or words in a sentence with more prominence than others. The prominent words are in a sense highlighted for the listener, and stand out from other non-prominent words. Your task is to mark words that you hear as prominent in this way."*

The 30 utterances were presented through headphones, one at a time, in the same order as the recall experiment. Participants listened to each utterance twice and selected the words heard as prominent, using a custom web interface developed by the author TM.

For each word in the test set we obtain a prominence score (p-score), which represents the proportion of participants who marked the word as prominent. We analyzed p-scores for the critical word with the varying AP (e.g. *University*), and for the word in the remaining portion of the subject NP (e.g. *of Kosovo*). In sentences with more than one content word following the critical word in the NP (e.g., *healthcare law* in [*The motivation for the healthcare law*]$_{NP}$…) we report the average p-score over all of them.

4.2. Results

To determine the effect of AP on the perceived prominence of the critical word and the content word(s) in the same NP, Poisson regressions were run on the p-scores (as counts) with AP as a predictor (treatment coding with AP2 (Early High, 1-2) as the reference level) in R [10]. Anovas were run with Anova() in the car package [15].

Table 3. Poisson Regression for the Critical Word
Intercept: AP2 (Early High, 1-2); Estimates are log-counts

Fixed Effects	Estimate	SE	z	p
Intercept	1.841	0.073	25.3	< .001
AP1 (2-1)	0.196	0.098	2	.046
AP3 (Unaccented)	-1.841	0.197	-9.37	< .001

Residual Deviance: 83.967 on 87 df
Type-III Sum of Squares for AP: $\chi^2(2) = 189.42$; $p < .001$

Table 4. Poisson Regression for following word(s) in NP
Intercept: AP2 (Early High, 1-2); Estimates are log-counts

Fixed Effects	Estimate	SE	z	p
Intercept	1.729	0.077	22.5	< .001
AP1 (2-1)	-0.174	0.114	-1.53	.126
AP3 (Unaccented)	0.330	0.101	3.27	.001

Residual Deviance: 71.985 on 87 df
Type-III Sum of Squares for AP: $\chi^2(2) = 24.59$; $p < .001$

We have two main results, illustrated here with the example sentence (1a). First, the critical word (*University*) is significantly more often perceived by our listeners as prominent when it has AP2 (Early High, 1-2), compared to AP3, (unaccented). However, the critical word is significantly less often perceived by our listeners as prominent when it has AP2 compared to AP1 (Table 3), though the difference between AP1 and AP2 (0.196) is much smaller than the difference between AP2 and AP3 (1.841). This yields the hierarchy AP1 > AP2 > AP3 for the critical word (Fig. 2, left panel).

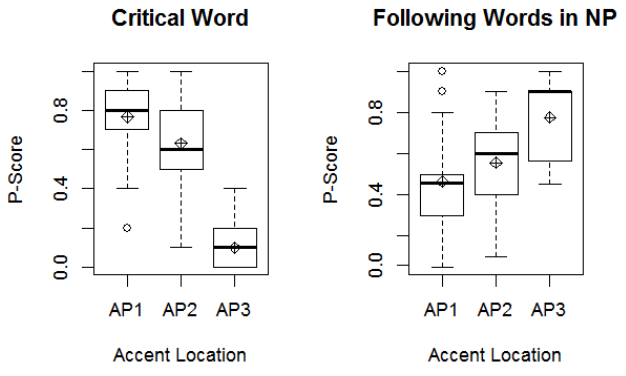
Conversely, the following content word(s) in the same NP (e.g. *Kosovo*) is (are) nearly always perceived as more prominent when the critical word (e.g. *University*) is unaccented. The early or late location of the accentual H on the critical word, on the other hand, does not significantly affect the perceived prominence of following words in the NP (Table 4). This yields the hierarchy AP3 > AP1, AP2 for the following content word(s) in the NP (Fig. 2, right panel).

These findings provide strong evidence that early high confers prominence on the accented word, as does the canonical pitch accent placement. In addition, the difference in p-scores between AP1 and AP2 suggests a less salient prominence for the early high pattern, which may reflect ambiguity in the function of early high as demarcative and/or prominence-lending. Finally, the finding that increased prominence on the critical word is associated with decreased prominence on the following word(s) in the subject NP is strong evidence for the relational character of stress/prominence at the phrasal level.

5. DISCUSSION

The results of our prominence labelling experiment show that the anchoring of the pitch accent on the syllable with lexical primary vs. secondary stress has significant but small effect on the degree of prominence of the accented word. This leaves open the question of why words with lexical 2-1 stress allow two stress patterns in contexts that do not involve stress clash. We hypothesized that the marked, early high pattern may be used to draw listeners' attention to the upcoming prosodic unit, perhaps marking the entire phrase as new information (or a new topic), as suggested by the frequent use of the early high pattern by broadcast news announcers. This hypothesis was tested in the recall experiment, but the results failed to show evidence that early high facilitates recall of sentence information beyond the accented word. It remains to be tested if early high facilitates recall locally, for the accented word, and how any such local effects of early high compare with canonical accenting.

Figure 2. P-scores of critical word (left) and of the following word(s) in the NP (right), grouped by AP.



6. ACKNOWLEDGMENTS

The authors thank Suyeon Im and Sarah Little for their assistance with this study, and Caroline Smith for helpful feedback. All errors remain with the authors. This study is supported by NSF BCS 12-51343.

7. REFERENCES

[1] Bolinger, D. 1981. Two kinds of vowels, to kinds of rhythm. Reproduced by the Indiana University Linguistics Club.

[2] Shattuck-Hufnagel, S. 1992. Early pitch accent placement within words in American English. *J. Acoust. Soc. Am.* 92(4), 2443.

[3] Shattuck-Hufnagel, S. 1995. The importance of phonological transcription in empirical approaches to 'stress shift' versus 'early accent'. In: Connell, B., Arvaniti, A., (eds), *Papers in Laboratory Phonology IV*, 128–140. Cambridge, UK: Cambridge University Press.

[4] Ross, K., Ostendorf, M., Shattuck-Hufnagel, S. 1992. Factors affecting pitch accent placement. *Proc. Of 1992 Int. Conf. On Spoken Lg. Processing*, Banff, 1, 365–368.

[5] Shattuck-Hufnagel, S. 1998. Acoustic-phonetic correlates of stress shift. *J. Acoust. Soc. Am.* 84(S1), S98.

[6] Beckman, M., Edwards, J. 1994. Articulatory evidence for differentiating stress categories. In: Keating, P., (ed), *Papers in Laboratory Phonology III*, 7–33. Cambridge, UK: Cambridge University Press.

[7] Grabe, E., Warren, P. 1995. Stress shift: do speakers do it or do listeners hear it. In: Connell, B., Arvaniti, A., (eds), *Papers in Laboratory Phonology IV*, 95–110. Cambridge, UK: Cambridge University Press.

[8] Fraundorf, S. H., Watson, D. G., Benjamin, A. S. 2010. Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *J. Mem. Lang.* 63(3), 367–386.

[9] Xu, Y. 2005. "Speech melody as articulatorily implemented communicative functions". *Speech Communication* 46, 220-251.

[10] R Core Team. 2014. R: A language and environment for statistical computing (Version 3.1.2). Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

[11] Bates, D., Maechler, M., Bolker, B. and Walker, S. 2014. lme4: Linear Mixed-Effects Models Using Eigen and S4. http://CRAN.R-project.org/package=lme4.

[12] Barr, D.J., Levy, R., Scheepers, C. and Tily, H.J. 2013. "Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal." *J. Mem. Lang.* 68 (3): 255–78.

[13] Singmann, H., Bolker, B. and Westfall, J. 2015. afex: Analysis of Factorial Experiments. http://CRAN.R-project.org/package=afex

[14] Cole, J., Mo, Y., Hasegawa-Johnson, M. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1, 425–452.

[15] Fox, J. and Weisberg, S. 2011. An R Companion to Applied Regression (version 2). Thousand Oaks, CA: Sage. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.