**Title:** Perception of stressed syllables varies by listener.
Amelia E. Kimball (akimbal2@illinois.edu), Jennifer Cole
University of Illinois at Urbana-Champaign

**Introduction:** The realization of stress in English phrases and sentences is theorized to involve stress at several levels of a prosodic hierarchy or metrical grid (e.g. Liberman & Prince, 1977). Word-level stresses are hypothesized to derive from metrical parsing algorithms (Hayes, 1995) that define a regular alternating pattern of stress prominence, avoiding two stresses next to one another (a clash) (Vogel, Bunnell, & Hoskins, 1995). At the sentence level, the word-level stresses concatenate, with assignment of stress based on syntactic structure (mediated through prosodic phrasing) and information structure (Cinque, 1993; Hammond, 1992). In a given utterance, stress production involves a complex interplay of the factors above.

Phrase-level stress patterns are important because regular patterns have been shown to affect processing of the linguistic signal in detection tasks (Quéné & Port, 2005; Zheng & Pierrrehumbert, 2010), as well as in production tasks (Tilsen, 2011). The pattern of stress over an utterance also has consequences for listeners in memory (Kimball, Yiu & Watson, submitted) and by extension, in comprehension. If this is the case, listeners should attend to the acoustic cues that express metrical patterns, analyze them, and—if all listeners are sensitive to the same constraints and patterns—arrive at the same perceived stress location.

However, several experiments indicate that listener’s perception of stress is not always predicted from available acoustic cues and that listeners may disagree with one another in the perceived location of stress. For example, while acoustic isochrony has been refuted (Dauer, 1983) perceptual isochrony is reported by annotators (Dilley, Wallace, & Heffner, 2012). Furthermore, naïve listeners disagree on the location of stressed syllables (Kimball & Cole, 2014), and even expert annotators may disagree on the occurrence of stress shift (Vogel et al., 1995) or location of stressed syllables (Temperley, 2008). Much research in stress perception is called into question, however, because stress reporting tasks are inherently metalinguistic. The observed lack of agreement, therefore, may be due to difference in skill at the metalinguistic task, or difference in interpretation of the task, rather than actual variability in stress perception. In light of these results, we ask: Is there variability in the perception of stress location, above and beyond variability due to task effects?

We use a metalinguistic stress reporting task similar to Kimball & Cole (2014) to measure this. We improve on previous studies by also including a post test that establishes listener skill, such that we can compare results from only those subjects who achieve high agreement in the post test. We predict that variability in stress perception exists, above and beyond the noise due to different metalinguistic skill levels.

**Procedure:** Participants were presented with an audio file, and a transcription of that file syllable by syllable, using Qualtrics survey software. Participants were told to “click on the checkboxes that you think correspond to the stressed syllables in the sentence.” 52 participants were recruited on Amazon Mechanical Turk (AMT), an online marketplace for human intelligence tasks. AMT screening tools ensured that the posting was displayed only to workers who reported their location as in the United States. Only data from the 48 participants who were self-reported English speakers with no hearing problems is reported.

**Stimuli:** The listeners heard three types of stimuli (see figure 1, below). First, they heard samples of sentences taken from the Buckeye corpus (Pitt et al., 2007) and pronounced by a trained linguist who was not a member of the research team. Next, they listened to individual words and were asked to mark the one syllable they thought was most stressed. Last, they listened to recordings of a limerick and the lyrics to Yankee Doodle, both of which involve a distinctive rhythmic pattern. The poem and individual words were spoken by the research team, and were produced in a manner that emphasizes the stress patterns.

<table>
<thead>
<tr>
<th>Buckeye</th>
<th>Words</th>
<th>Poem</th>
</tr>
</thead>
<tbody>
<tr>
<td>I think of it as a big city</td>
<td>certificate</td>
<td>Yankee Doodle went to town.</td>
</tr>
</tbody>
</table>

Fig. 1: Example words from each stimuli set.

**Results and analysis:** Each syllable received a rating based on the number of listeners that marked it as a beat. Agreement scores are proportions ranging from 0 to 1. If listeners were in total agreement, the distribution of these scores would be bimodal, with peaks at 0 (meaning many syllables were marked by no listeners) and 1 (meaning many syllables were marked by all listeners). We compare the agreement scores
in three groups: Our original result from our 2014 experiment, all participants from this experiment, and just those participants who scored above 80% on the words and poem combined (15 out of 48 participants). A Mann-Whitney test showed that subjects who scored higher than 80% on the post-test had a different distribution than both those who scored under 80% ($W= 26827$, $p<.001$) and subjects in our original experiment ($W= 25520$, $p<.001$). Crucially, though the groups differed, none of the groups achieved a bimodal distribution indicating high agreement.

**Discussion:** Our results indicate a lack of agreement on the location of stressed syllables, even among those participants who have high agreement for individual words or patterned speech such as poems. Based on these results, we argue that the realization of phrasal stress is complex not only in the varying constraints that underlie the acoustic realization, but also in listener-to-listener variation in perception. We hypothesize that top-down processing based on language learning over the lifetime and skills such as musical training affect perception of stress. Overall, we suggest that in order to understand phrasal stress, listener perception must be modeled in addition to modeling metrical structure and the acoustic spell-out of stress.

**References:**


Kimball, A.E., Yiu, L.K., & Watson, D. (Under review) *Word recall is affected by surrounding metrical context.*


