Nathaniel B Clark and Marcus Perlman, University of California, Santa Cruz

Visualizing vocal iconicity: Empirical and graphical methods for investigating pitch contours

While iconicity is becoming more widely acknowledged as commonplace, even in spoken language (Perniss, et al., 2010), the empirical study of iconicity in speech is hampered by a reliance on comparing averages for acoustic variables across whole words or phrases. The collapse of continuous prosodic information into averages greatly reduces the level of detail that can be examined in studies of vocal iconicity (e.g., Shintel, et al., 2006.) This paper presents a novel methodological approach, borrowing ideas from mousetracking, to explore the iconic potential of pitch and other continuous prosodic variables. We summarize the method with a graphical analysis of pilot data, and discuss two statistical methods which will be applied to data from an ongoing study. Finally, we discuss some implications and future uses of this work.

The pilot results we present here are derived from an earlier study (Perlman, et al., under review), in which participants read short stories discussing physical and abstract instantiations of UP and DOWN. We focused on the phrase "all the way," which appeared in two stories describing vertically-oriented physical motion. One story described Chris riding an elevator up (down) to the top floor (basement) of a skyscraper, and the other described Melinda reaching up (bending down) to get an item from the very top (bottom) shelf at a store. The phrase was identical across conditions; only context varied. Thus, differences between pitch contours should be due to semantic context rather than phonological content. Each of the 30 participants recruited for this study read both of these stories, 15 with Chris in the UP condition and 15 with Melinda in the UP condition.

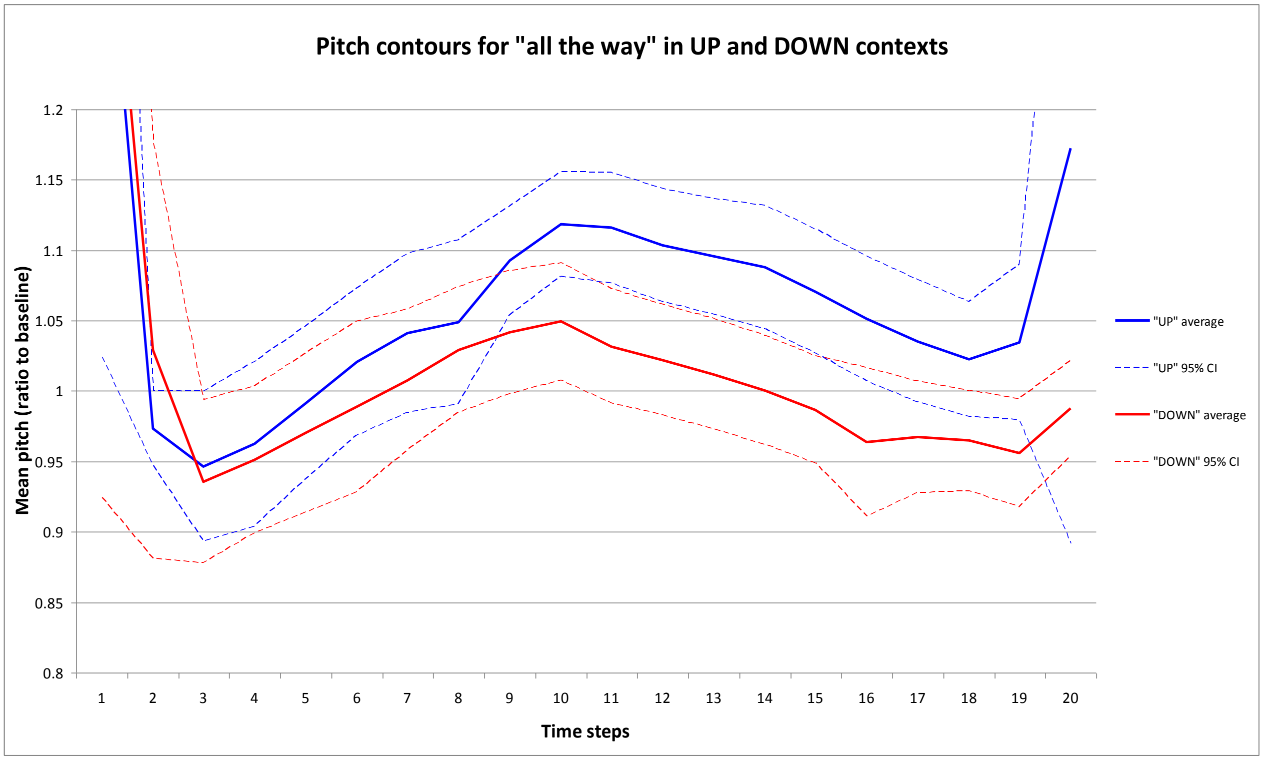
We segmented the phrases in Praat (Boersma and Weenick, 2011). A script sliced each phrase into 20 intervals and computed the average pitch in each ventile. For the graphical analysis, we normalized for individual differences in pitch by dividing each observation by that speakers' baseline pitch, calculated from control utterances.

Figure 1 illustrates the results of this analysis. The very large confidence intervals at the left and right extrema indicate problems with missing data due to pitch tracker inaccuracies and amodal voicing. The critical section, though, is the 11th through 15th intervals, in which the CIs for UP are wholly above those for DOWN, indicating that participants exhibit an iconically motivated difference in pitch of about 8% (for all 5, $t(29) > 2.19$, $p < .05$).

Based on these promising results, we are implementing a new set of stimuli designed specifically for this methodology. Our first statistical method is a bootstrap analysis (cf. Dale, et al., 2007), to identify a significance criterion for consecutive outcomes and reduce the risk of Type-I error. Second, growth curve analysis (cf. Mirman, et al., 2008) will be used to probe for significant differences in change in pitch over time. This method takes into account the data's temporal sequence, and has the additional advantage of obviating the need for normalization across speakers, since the hierarchical nature of the model is well-suited for individual differences.

While vocal iconicity has been notoriously intractable to empirical study (e.g., McNeill, 2005), these mouse-tracking derived approaches offer powerful tools for future examination. By affording two-dimensional visualizations of acoustic variables, they permit more fine-grained analysis of continuous speech, and ultimately allow for more sophisticated comparisons to visual iconic gesture.

Figure 1.



References

Boersma, P. and Weenick, D. *Praat: Doing phonetics by computer* [computer program]. Version 5.3.03, retrieved 28 November 2011 from <http://www.praat.org/>.

Dale, R. Kehoe, C., and Spivey, M. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory and Cognition*, *35*, 15-28.

Mirman, D. Dixon, J.A., & Magnuson, J.S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59(4), 475-494.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

Perlman, M., Clark, N., and Johansson-Falck, M. (under review). Iconic modulations of prosody during story reading. [Manuscript under review].

Perniss, R., Thompson, R., and Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*(227).

Shintel, H., Nusbaum, H., and Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, *55*, 167-177.