

The combined influence of intensity and duration on rhythmic grouping in speech

Rhythm in language is conveyed by the alternating distribution of stressed and unstressed syllables in patterns that phonologists interpret as evidence for *prosodic feet*, recurrent categories that are posited as elements of grammar. Decades of research in phonology have produced extensive knowledge of the forms feet can take and their cross-linguistic patterning (e.g. Hayes 1985, 1987, 1995; McCarthy & Prince 1986). A widely accepted foot typology proposed by Hayes (1987, 1995) includes trochaic (left-dominant) and iambic (right-dominant) feet. A complementary tradition of phonetic research has identified acoustic correlates of stressed syllables and their influence on listeners' stress judgments (e.g. Fry 1958, 1965; Lehiste 1970; Beckman 1986; Beckman & Edwards 1994; Sluijter & van Heuven. 1996; Sluijter *et al* 1997). However, we still know relatively little about the perceptual factors that shape the groupings themselves, as reflected in listeners' biases when they are asked to subjectively group rhythmically alternating polysyllabic speech sequences. Early psychoacoustic research on rhythm perception has revealed two grouping principles, which Hayes cites in support of his foot typology (Hayes 1995:80, following Bolton 1894, Woodward 1909):

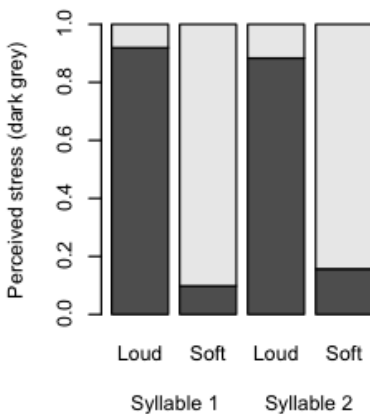
- (i) elements contrasting in intensity naturally form groupings with initial prominence.
- (ii) elements contrasting in duration naturally form groupings with final prominence.

Modern work on grouping rhythmic *nonspeech* sequences has found varying support for these principles (Trainor & Adams 2000, Kusumoto & Moreton 1997, Hay & Diehl 2007, Iversen & Patel 2007). Hay & Diehl (2007) report that principles (i) and (ii) also predict English-speaking and French-speaking listeners' subjective grouping of alternating *speech*-like sequences. Hay & Diehl's findings have limited applicability to speech, however, because in speech, vowel duration and intensity vary together and interact in intricate ways (Turk & Sawusch 1996).

This paper reports the findings of two experiments exploring the joint influence of duration and intensity on the subjective grouping of rhythmic speech sequences. In Experiment 1, listeners heard bisyllables in which vowel intensity and duration were both varied, synchronously (loud/long + soft/short, or the reverse), and asynchronously (loud/short + soft/long, or the reverse). Asked to indicate which syllable was stressed, listeners overwhelmingly perceived the louder syllable as stressed, regardless of its length or position (Figure 1). Experiment 2 was conducted with the same subjects and during the same session as Experiment 1. Subjects heard longer alternating strings keyed to the pairs constructed for Experiment 1. They were instructed to indicate whether the syllable they heard as stressed came first or last in the repeating pairs. The results of Experiment 2 reveal that while vowel intensity was the more robust predictor of stress judgments, duration was the better predictor of grouping (Figure 2).

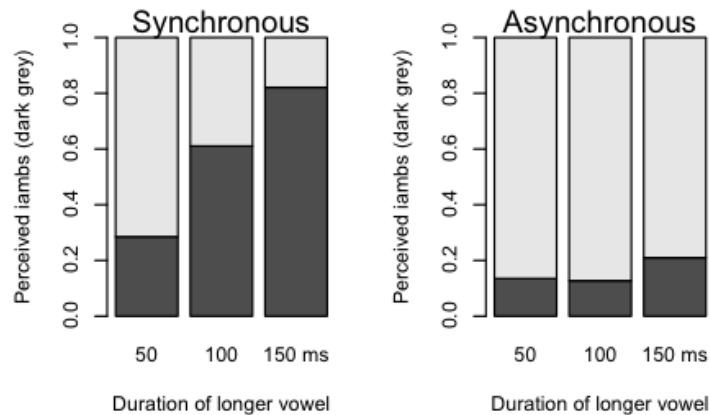
Ultimately, if we can provide a comprehensive account of listeners' grouping biases, identify the perceptual factors that shape them, isolate these from any role they may also have in cuing stress, and eventually, study the extent to which these biases are general across speakers of different languages – then we will have important experimental support for the rhythmic groupings underlying the foot inventory that phonologists have established based primarily on distributional evidence.

Experiment 1: Bisyllables.



The stressed syllable is perceived to be the louder of the pair, regardless of position (1st or 2nd syllable).

Experiment 2: Polysyllabic sequences.



Proportion of perceived iambs in Experiment 3. *Asynchronous*: When intensity is separated from length, odds of an iamb grouping are low. (Loud is stressed; grouping principles (i) and (ii) are satisfied.) *Synchronous condition*: when the same syllable is both long and loud, odds of an iamb decision rise in proportion to increases in duration. (Grouping principle (ii) trumps grouping principle (i).)

References

- Beckman, Mary. 1986. *Stress and non-stress accent*. Dordrecht: Foris Publications.
- Beckman, Mary E. and J. Edwards. 1994. 'Articulatory evidence for differentiating stress categories', in P. A. Keating (eds.) *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press, pp. 1–33.
- Bolton, Thaddeus. 1894. 'Rhythm', *American Journal of Psychology* 6, 145-238.
- Fry, D. B. 1958. 'Experiments in the perception of stress', *Language and Speech* 1, 126–152.
- Fry, D. B. 1965. 'The dependence of stress judgments on vowel formant structure', *Proceedings of the 5th International Congress on Phonetic Science, Münster, 1964* (Karger, Basel), pp. 306–311.
- Hay, Jessica & Randy Diehl. 2007. 'Perception of rhythmic grouping: Testing the iambic/trochaic law', *Perception & Psychophysics* 69, 113-122.
- Hayes, Bruce. 1985. *A metrical Theory of Stress Rules*. NY: Garland Press.
- Hayes, Bruce. 1987. 'A revised parametric metrical theory', *Northeastern Linguistics Society* 17, 274-289.
- Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
- Iversen, John R. & Aniruddh D. Patel. 2008. 'Perception of rhythmic grouping depends on auditory experience', *JASA* 124, 2263–2271.
- Kusumoto, K., Moreton, E. 1997. 'Native language determines parsing of nonlinguistic rhythmic stimuli' *JASA* 105, 3204.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: The MIT Press.
- John J. McCarthy and Alan Prince. *Prosodic Morphology 1986*. Technical Report #32, Rutgers University Center for Cognitive Science, 1996.
- Sluijter, Agaath M. C. & Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *JASA*, 2471-2485.
- Sluijter, Agaath M. C., Vincent J. van Heuven, & Jos J. A. Pacilly. 1997. Spectral balance as a cue in the perception of linguistic stress. *JASA* 101, 503-513.
- Trainor, Laurel & Beth Adams. 2000. Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception and psychophysics* 62, 333-340.
- Turk, Alice & J. Sawusch. 1996. 'The processing of duration and intensity cues to prominence,' *JASA* 99, 3782–3790.
- Woodrow, Herbert. 1909. *A Quantitative Study of Rhythm: The effect of variations in intensity, rate, and duration*. Columbia University Contributions to Philosophy and Psychology, Vol. XVIII, No. 1.