Pitch trumps duration in a grouping perception task

Alejna Brugos & Jonathan Barnes (Boston University)
abrugos@bu.edu

Prosody; Perception; Boundary perception; Phrasing; Prosodic grouping; Pitch; Timing; American English

Speech timing patterns are known to serve as critical cues for perceived prosodic grouping, making timing central to the study of how prosody encodes meaning at all levels of linguistic structure. Investigations of, e.g., the prosody of attachment ambiguities therefore focus heavily on temporal information, operationalized as objective interval duration. Experimental studies of F0 in the context of prosodic grouping are comparatively rare (cf. [7], [4] on "declination reset", [8] on phrase-initial reset). Perceived duration, however, may differ dramatically from measured duration: Dynamic f0 in speech can lead to longer perceived vowel duration ([12], [3]), and non-speech research has likewise shown that pitch manipulations can alter perception of timing ([9], [6]). Studies on the auditory kappa effect in particular ([2], [5]), show that in sequences of tones and silent intervals, pitch differences among tones can distort perception of timing such that tones closer in pitch are also perceived as closer in time.

To determine whether pitch interacts with timing similarly in the perception of prosodic grouping in speech, we conducted a study involving a string of 3 spoken numbers, parsable as "NN-N" or "N-NN". The AXB design (based on [10] and [9]), used 3 rise-fall, full intonational phrase (H* L-L%) versions of the word *one,* resynthesized from the same 302 ms. base recording, and shifted in 1-semitone steps. A was set as the highest, 8 semitones above B. X was chosen from 7 intermediate pitch steps, and placed at each of 10 time steps (410 to 590 ms.) after A. (The X to B interval likewise shifted, such that the two silent intervals always totalled 1 second.) 14 participants indicated whether X was grouped with A or B for 4 repetitions of 70 resulting stimuli.

A previous study [1] with identical stimuli showed that, in a non-linguistic task involving explicit timing comparison of the A-to-X and X-to-B intervals, while subject responses were based primarily on interval duration, they were modulated by relative pitch as well: As with the kappa effect in non-speech studies, closer in pitch sounded closer in time. In the present study, however, subjects received no instructions concerning the signal itself, but were told to report only which "grouping" they heard. Surprisingly, timing now affected subject responses comparatively little: where X was closest to A (e.g., 1 or 2 st. lower), subjects overwhelmingly grouped X with A, while X closer to B (e.g., 6 or 7 st. below A) cued grouping with B, interval durations notwithstanding. Timing differences affected responses strongly only for intermediate (i.e. ambiguous) pitch steps. (See partial summary in Table 1.)

These results suggest that quantification of boundary strength based only on objective duration misses powerful cues from F0. This may shed light on durational variability in related production studies: jumps in pitch across pauses may signal stronger boundaries, while steady pitch may signal a weaker boundary, in ways that current systems of categorical pitch event labels (e.g., ToBI) are not designed to capture. These results also parallel findings from duration studies characterizing boundary strength as inherently relative, and gradiently variable [11].

Table 1: Partial Results	**time** \ **pitch**	**1 st.**	**2 st.**	**4 st.**	**6 st.**	**7 st.**
% responses of	A to X = 410 ms. (B to X = 590 ms.)	4%	7%	25%	69%	78%
"X grouped with B" for a	A to X = 470 ms. (B to X = 530 ms.)	2%	11%	49%	83%	87%
subset of time & pitch	A to X = 530 ms. (B to X = 470 ms.)	11%	11%	43%	89%	95%
steps	A to X = 590 ms. (B to X = 410 ms.)	17%	29%	61%	84%	96%

References

[1] Brugos & Barnes (2012) The auditory kappa effect in a speech context. *Speech Prosody*, Shanghai.
[2] Cohen, Hansel, & Sylvester (1954) Interdependence of temporal and auditory...*Nature*,*174*: 642–644.
[3] Cumming (2011) … dynamic fundamental frequency on the perception of duration, *JPhon*, 39(3): 375–387.
[4] Féry & Truckenbrodt (2005) Sisterhood and tonal scaling, *Studia Linguistica*, 59(3): 223-243.
[5] Henry & McAuley (2009) …imputed pitch velocity model…auditory kappa effect, *J. Exp Psych,35*(2): 551–564.
[6] Henry (2011) *A Test of an Auditory Motion Hypothesis for…Sounds Moving in Pitch Space*. Thesis. B.G.S.U.
[7] Ladd (1988) Declination 'reset' and the hierarchical organization of utterances. *JASA*, 84: 530-544.
[8] Lin & Fon (2011) The role of pitch reset in perception at discourse boundaries, *ICPhS XVII*, Hong Kong.
[9] MacKenzie (2007) *The kappa effect in pitch/time context*. Thesis. Ohio State University.
[10] Shigeno (1986) The auditory tau and kappa effects…, *Percep & Psychophysics*, 40(1): 9–19.
[11] Wagner & Crivellaro (2010) Relative Prosodic Boundary Strength…, *Speech Prosody*, Chicago
[12] Yu (2010) Tonal effects on perceived vowel duration. Fougeron, et al (eds.), *LabPhon 10*. Mouton de Gruyter.