Pitch trumps duration in a grouping perception task

Alejna Brugos and Jonathan Barnes (Boston University)
abrugos@bu.edu, jabarnes@bu.edu

Prosodic grouping: Cues from timing and pitch

The phonetic realization of prosodic grouping includes aspects of timing, pitch, as well as segmental, voice quality and amplitude cues.

Speech timing patterns are critical cues for perceived prosodic grouping

(Wightman et al. 1992)

- Central to the study of how prosody encodes meaning at all levels of linguistic structure
- Operationalized as objective interval duration (of, e.g., segments, syllables, silent pauses)

F0 cues are also recognized as important to grouping

- Pitch accent scaling (Ladd, 1988; Féry & Truckenbrodt, 2005)
- Phrase-initial reset (Jun, 2006; Lin & Fon, 2011)
- Phrase accents and boundary tones (Beckman & Ayers Elam,1997)

There is some evidence that f0 cues may contribute less than those of timing

(Holzgrefe et al 2011, Hansson, 2003)

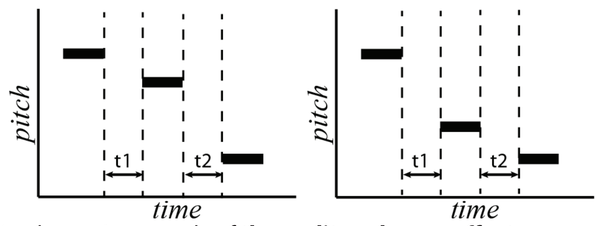
--> f0 cues left out of some grouping studies (Wagner & Crivellaro, 2010, Holsinger et al. 2010)

However, the perception of pitch and time may not be independent

Perceived duration may differ dramatically from measured duration:

- Dynamic f0 in speech can lead to longer perceived vowel duration (Yu, 2010; Cumming, 2011)
- Non-speech research showing that pitch manipulations can alter perception of timing (Crowder & Neath, 1995 ; Henry, 2011)
  - The auditory kappa effect (Cohen et al., 1954; Henry & McAuley, 2009; *inter alia*)

The auditory kappa effect:
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.

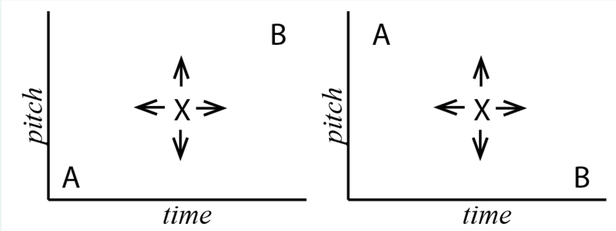


*A schematic example of **the auditory kappa effect**: The silent intervals (t1 and t2) are of equal duration, but t1 is perceived as shorter at left, longer at right.*

Investigating F0/time interaction in a speech context

This study: 2 new experiments

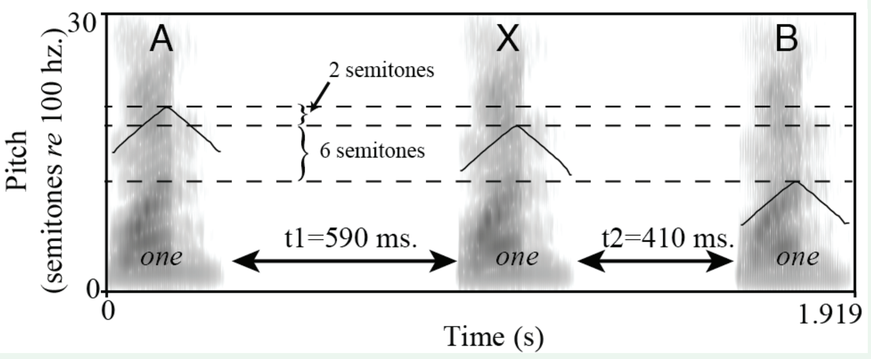
- Modelled after studies on auditory kappa effect
- Used AXB kappa cell paradigm (Shigeno, 1986; MacKenzie, 2007)
  - Sound events A and B fixed in pitch space, and in time relative to each other
  - Only intermediate event X changes, in both time and pitch space.



*A schematic of **the kappa cell paradigm** for testing the auditory kappa effect for two pitch change directions: ascending (left) and descending (right).*

- Using speech stimuli (See also Brugos & Barnes, 2012):
  - String of 3 spoken numbers, parsable as "NN-N" or "N-NN"
    - Single-word full intonational phrase (H* L-L%) resynthesized versions of the word *one*
      - From the same 302 ms. base recording, shifted in 1-semitone steps
    - Concatenated in 2 pitch change dirctions, descending and ascending
    - A set as the highest (or lowest), 8 semitones above (or below) B
    - X at 7 intermediate pitch steps
      - Placed at each of 10 time steps (410 to 590 ms.) after A
      - The X to B interval likewise shifted --> 2 silent intervals always totalled 1000 ms.
    - 4 repetitions of 70 resulting stimuli

***Sample stimulus file**:*
A sequence of the spoken word one, *from the descending direction condition. The f0 contour of the first* one *(A) is 8 semitones above the f0 contour of the third* one *(B). The middle* one *(X) is set to 2 st. below A. The silent interval between A and X is set to 590 ms.*



Identical stimuli used in both experiments: Only the tasks differ

Experiment 1: The auditory kappa effect in a speech context

Methods: Subjects indicated whether X was closer *in time* to A or B

- explicitly directed to ignore pitch
- 2 orders (ascending & descending), 2 groups of subjects (N=31)

Results: Subject responses based primarily on interval duration, but modulated by relative pitch. As with the kappa effect in non-speech studies, closer in pitch sounded closer in time.

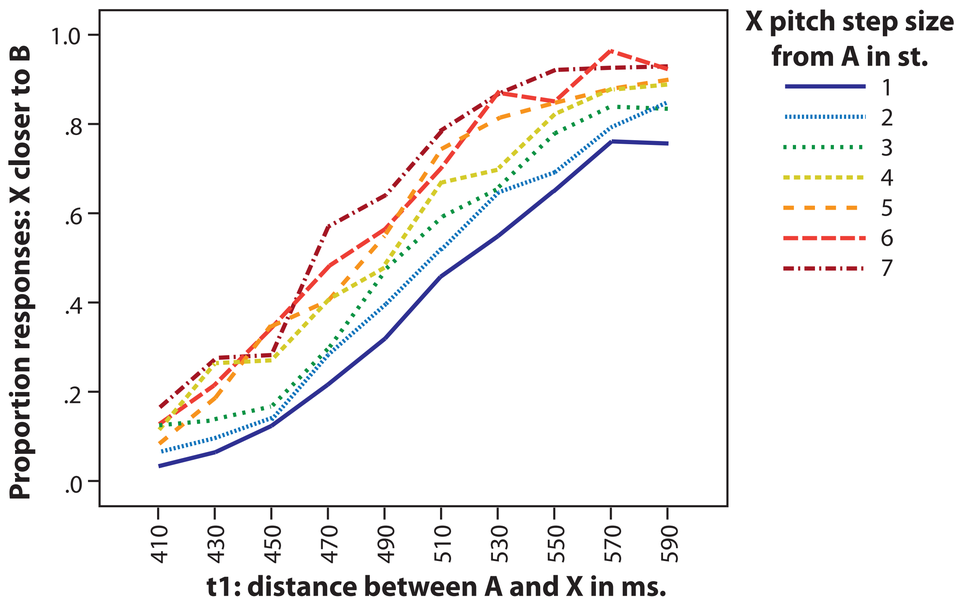
Experiment 2: Perception of grouping

Methods: Subjects indicated whether X was "grouped" with A or B

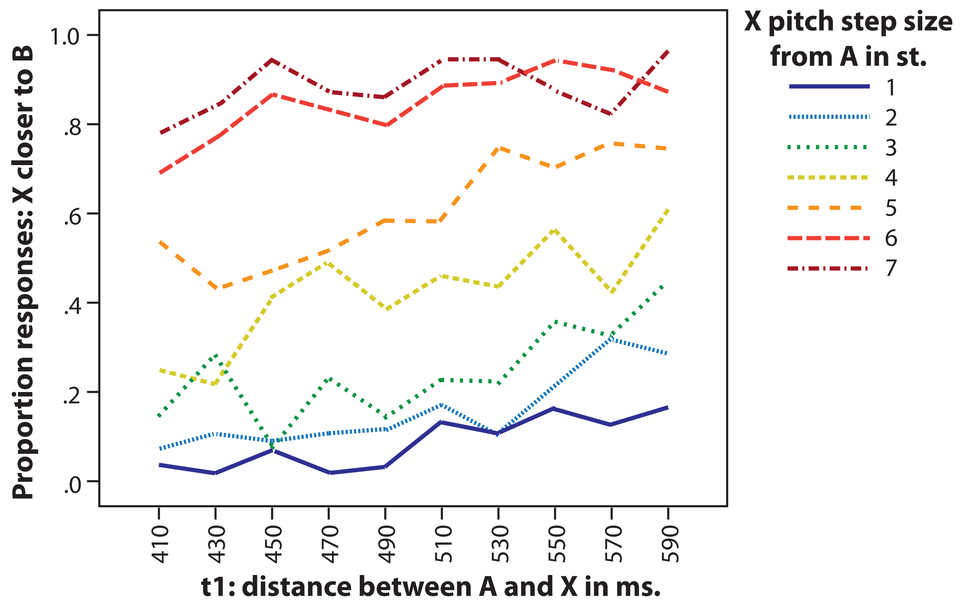
- no instructions concerning the signal itself
- stimuli from descending order (1 group, N=14)

Results: When X was closer to A in pitch, subjects grouped X with A, closer to B cued grouping with B. Timing affected responses strongly only for intermediate (i.e. ambiguous) pitch steps.

Results of experiment 1: Timing perception modulated by pitch



Results of experiment 2: Grouping perception dominated by pitch



Conclusions: Quantification of boundary strength based only on objective duration misses powerful cues from F0.

Results may shed light on mismatches of durational patterns and phrasing perception

- Jumps in pitch across pauses may signal stronger boundaries
- Steady pitch may signal a weaker boundary

--> Pitch and timing may be in a cue trading relationship (Beach, 1991)

Results also parallel findings from duration studies characterizing boundary strength as inherently relative, and gradiently variable (Wagner & Crivellaro, 2010)

- Pitch change across phrases may be gradiently implemented
- May reflect cross-IP patterns that current systems of categorical pitch event labels (e.g., ToBI) are not designed to capture.

Acknowledgments: This work was supported by NSF grant #1023853.

References:

Beach, C. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, 30(6): 644–663.
Beckman, M., & Ayers Elam, G. (1997). Guidelines for ToBI Labelling. (March, 1997,v. 3).
Brugos, A. & Barnes, J. (2012). The auditory kappa effect in a speech context. *SpPros,* Shanghai.
Cohen, J., Hansel, C., & Sylvester, J. (1953). A new phenomenon in time judgment, *Nature*, 172: 901.
Cohen, J., Hansel, C. & Sylvester, J. (1954). Interdependence of temporal and auditory judgments. *Nature*, 174: 642–644.
Crowder, R. & Neath, I. (1995). The influence of pitch on time perception in short melodies. *Music Perception*, 12(4): 379–386.
Cumming, R. (2001). The effect of dynamic fundamental frequency on the perception of duration. *JPhon*, 39(3): 375–387.
Féry, C. & Truckenbrodt, H. (2005). Sisterhood and tonal scaling. *Studia Linguistica*, 59(3): 223-243.
Henry, M. & McAuley, J. (2009). Evaluation of an imputed pitch velocity model of the auditory kappa effect. *Journal of Experimental Psychology: HPP*, 35(2): 551–564.
Henry, M. (2011). *A Test of an Auditory Motion Hypothesis for Continous and Discrete Sounds Moving in Pitch Space*. PhD. Dissertation. Bowling Green State University.
Holsinger, E., Li, D., Kaiser, E. & Byrd, D. (2010). Visual Grouping and Prosodic Grouing: Effects of Spatial Information on Prosodic Boundary Strength. *SpPros*, Chicago.
Holzgrefe, J., Schröder, C., Höhle, B. & Wartenburger, I. (2011). Neurophysiological investigations on the processing of prosodic boundary cues. *ETAP 2*, Montreal.
Jun, S.-A. (2006). Intonational phonology of Seoul Korean revisited. In T. Vance & K. Jones (Eds.), *Japanese/Korean Linguistics* 14 (p. 15-26). Stanford: CSLI.
Ladd, D. (1988). Declination 'reset' and the hierarchical organization of utterances. *JASA*, 84: 530-544.
Lin, H. & Fon, J. (2011). The role of pitch reset in perception at discourse boundaries. *ICPhS* XVII, Hong Kong.
MacKenzie, N. (2007). *The kappa effect in pitch/time context*. PhD. Dissertation, OSU.
Shigeno, S. (1986). The auditory tau and kappa effects for speech and nonspeech stimuli. *Perception & Psychophysics*, 40(1): 9–19.
Wagner, M. & Crivellaro, (2010). Relative Prosodic Boundary Strength and Prior Bias in Disambiguation. *SpPros*, Chicago.
Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *JASA*, 91: 1707–1717.
Yu, A. (2010). Tonal effects on perceived vowel duration. In C. Fougeron, B. Kühnert, M. D'Imperio & N. Vallée (Eds.), *Papers in Lab. Phon.* (10). Berlin: M. de Gruyter.