

# Gradient perception of the stop voicing contrast in English and Korean: Evidence from eye-tracking

Eun Jong Kong, Jan Edwards University of Wisconsin-Madison

## BACKGROUND

### Two issues:

#### 1. Cross-linguistic differences in sensitivity to acoustic cues

- English: two-way contrast between voiced (unaspirated) and voiceless (aspirated) stops
- VOT alone can differentiate the contrast.
- fo is a secondary cue.
- Korean: three-way contrast among lax, tense, and aspirated stops.
- Both VOT and fo are necessary cues.
- Prediction: Korean listeners will be more sensitive to fo cues (in English) than English listeners.

#### 2. Individual differences in speech perception

- Auditory vs. phonemic listeners: Are some listeners more *categorical* and others more *gradient* in speech perception tasks?
- Prediction: Gradient listeners will be more sensitive than categorical listeners to subphonemic fine phonetic detail.

### Focus of this study

→ We investigated perception of the stop voicing contrast as a function of two acoustic cues, VOT and fo.

→ Three questions were addressed:

- Are there cross-linguistic differences in the use of fo as a cue to the stop voicing contrast between L1 English and L2 English speakers (L1 = Korean)?
- Among English speakers, are some listeners more categorical and some more gradient in their perception of the stop voicing contrast?
- If so, are the gradient listeners more sensitive to the secondary fo cue to voicing?

## METHODS

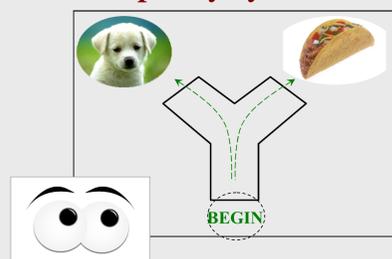
### Materials:

- Synthesized CV syllables based on words (*tot* and *dot*) produced by a Wisconsin adult male speaker.
- VOT values manipulated by excising a portion of the burst release/aspiration from /ta/ and pasting it before the voicing onset of the /da/ token.
- 30 different stimuli (6-step VOT × 5-step fo):
  - VOT steps: 9ms (original VOT of /da/), 13ms, 19ms, 28ms, 40ms, 59ms.
  - fo steps: 98Hz, 106Hz, 114Hz, 122Hz, 130Hz.

### Participants:

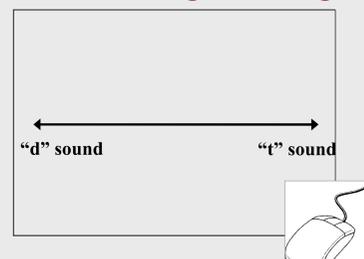
- 24 English- & 12 Korean-speaking adults (length of stay at US > 5 years)

### Tasks: Anticipatory Eye Movement



**Analysis (AEM):** Looking to /d/ target as a function of time according to each acoustic condition (time-series analysis in a mixed-effects regression model, Barr 2008, Mirman et al., 2008)

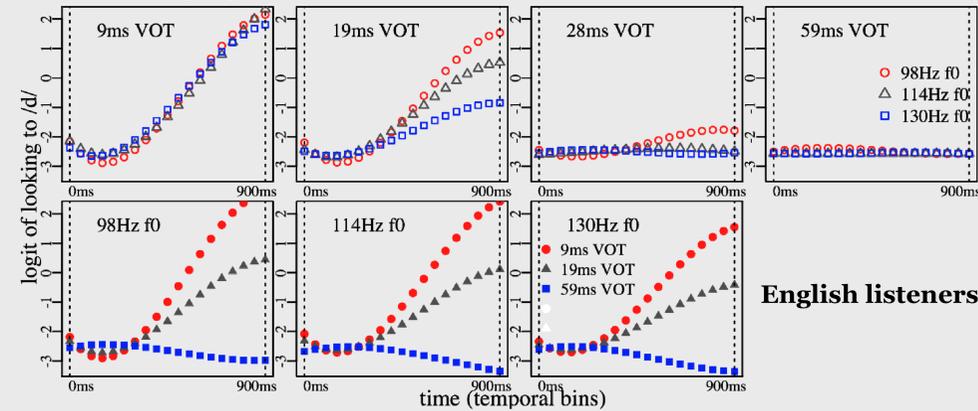
### Visual Analogue Scaling



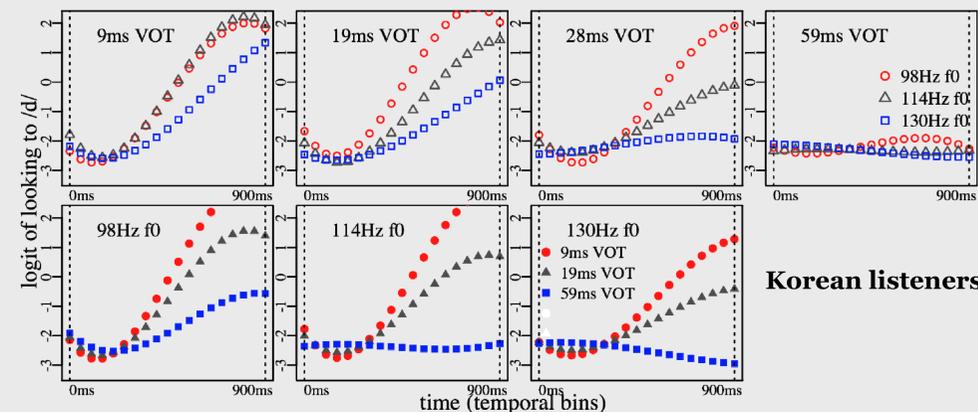
**Analysis (VAS):** Click location distributions

## RESULTS

### 1. Cross language differences

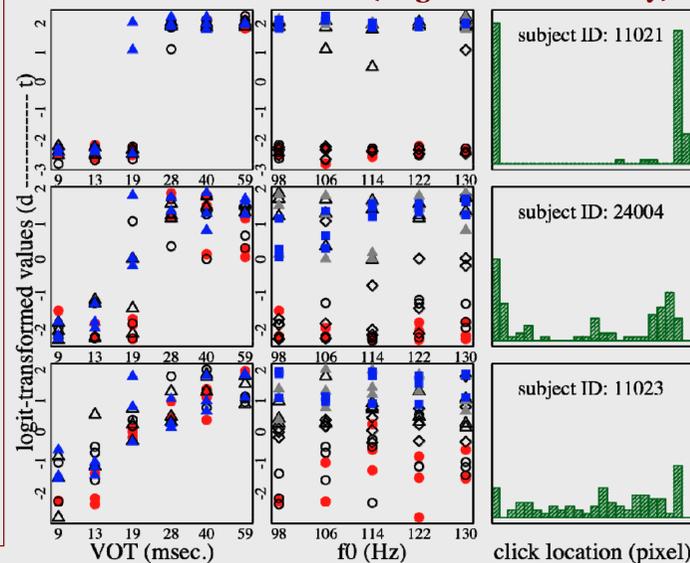


- Top row: little variation as a function of differences in fo, except 19ms VOT condition.
  - Only in ambiguous VOT (19 ms) condition, higher fo had a slower speed of looking to 'd'.
- Bottom row: Regardless of fo condition, more looking to 'd' in shorter VOT condition.



- Top row: Consistent effect of fo on looking to 'd' in all VOT conditions except 59ms.
- Bottom row: When fo is low (left-most plot), model fit line for long VOT condition shows influence of both cues (compare to similar plot for English listeners).

### 2. Individual differences (English listeners only)



### Visual Analogue Scaling Task

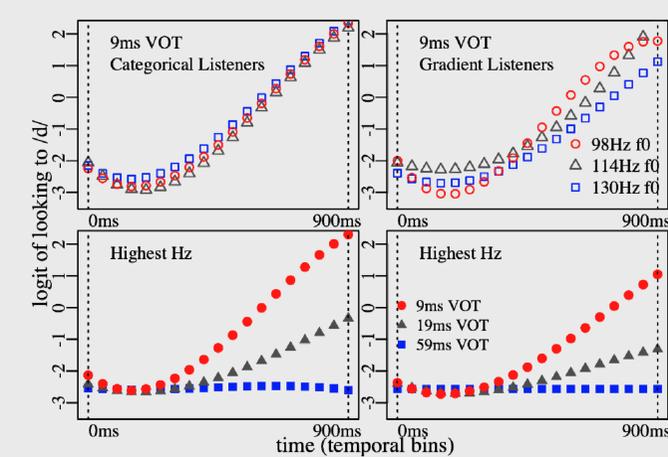
Three different patterns of click location distribution:

**Two-peak type:** 6 listeners judged the stimuli *categorically*, by choosing responses mostly at the two endpoints of /d/ and /t/.

**No-peak type:** 7 listeners judged the stimuli in a much more *gradient* manner, by choosing responses across the VAS scale.

**Intermediate type:** 11 listeners did not fit clearly into either group based on click location histograms.

### Anticipatory Eye Movement Task



### Highest fo condition:

- The slopes of the **categorical** (two-peak) group were influenced only by VOT.
- The slopes of the **gradient** (no-peak) group showed an influence of fo.
  - Shallower slopes relative to the categorical group for the two shorter VOT conditions (fo cue is in conflict with VOT cue).

## DISCUSSION & CONCLUSION

### Question 1: Cross-linguistic differences

- Cross language differences were found in listeners' sensitivity to fo cue
- Korean listeners sensitive to both fo and VOT for perception of stop voicing contrast.
- English listeners rely on VOT only.

### Question 2: Individual differences among English listeners

- Individual differences were also found in the perception of within-category differences:
  - About 25% of our listeners perceived the stimuli categorically, even on a VAS task that was designed to encourage gradient perception.
  - Another 25% of our listeners had a gradient pattern of response on the VAS task.

### Question 3: Gradient listeners and attention to fine phonetic detail

- Only the gradient listener group was sensitive to changes in fo on the AEM task.

Further research is needed to understand individual differences:

- First, it needs to be confirmed that these individual differences in speech perception are consistent.
- Then, we need to investigate whether there are subject-level characteristics that are consistently associated with these individual differences.

To conclude, this study is important in two respects:

- First, it illustrates how an online measure such as eye-tracking can provide information about aspects of speech perception that cannot be detected by offline measures.
- Second, it provides evidence that there are differences in how sensitive individuals are to fine phonetic detail and suggests that these differences may be related to different patterns of attention to acoustic cues.

## ACKNOWLEDGEMENTS

This research was supported by NIDCD Grant 02932 and NSF grant BCS 0729140 to Jan Edwards, and NSF Grant BCS 0739206 to Mary E. Beckman.