Introduction

- **Incomplete Neutralization (IN)**
  - /X/ $\rightarrow$ [Z]/ (Context A)
  - /Y/ $\rightarrow$ [Z]/ (Context A)
  - Final devoicing is a commonly cited case of IN.
- **AmE Flapping** is a potential case of IN:
  - /t/ and /d/ become [ç] in certain prosodic configurations (Kahn 1980), but distinctions remain on the surface.
  - Most notably, vowels before /d/-flaps are longer than vowels before /t/-flaps (Braver 2011, Herd et al. 2010) (though see Port (1976)).
  - This distinction is generally less than 10ms.

Stimuli

- Tokens were trisyllabic nonce words taken from a related production experiment (Braver 2011).
- 12 speakers produced each token in two tasks (no significant differences were found across tasks):
  - $\sigma$ 1: onsets {p,t,b,d}:
    - $\sigma$ 2: onsets {p,t,k}:
    - Nuclei: {i, e, ae}:
    - Codas: {d,t}
  - $\sigma$ 3: -ing (places d/t in flapping environment)
- Some representative minimal pairs:
  - puhPEEting ing puhPEEding
  - tuhKAT-ing tuhKAT-ding
  - duhTEHT-ing duhTEHD-ing
- Tokens were selected from three speakers, based on the following criteria:
  - Largest difference between pre-/d/ and pre-/t/ vowel duration.
  - Accurate production of a sufficient number of tokens.
  - Balanced on set and vowel of target syllable, as well as /d/ vs. /t/.

Methods, Part I

- 42 undergraduates participated in two tasks (21 per task).
- Feedback was given on each trial.
- Three blocks, each from a different speaker.
- Block order was balanced (Latin Square) across all listeners.
- **Identification Task:**
  - Listeners heard a single token, and were asked whether the sound before the -ing was a /d/ or a /t/.
- **ABX Task:**
  - Listeners heard three stimuli per trial (A, B, then, X), and were asked to decide whether X was the same as A or B.

Results, Part I (Continued)

- Hits vs. False Alarms for the Identification Task:
  - Listeners’ d’ scores were significantly different from 0 (mean $d' = -1.20$, Wilcoxon test: $V = 231$, $p < 0.001$).
  - Listeners said “X is like A” more often when X was actually like A than when X was actually like B.

- **ABX Task:**
  - Listeners’ d’ scores were not significantly different from 0 (mean $d' = -0.02$, Wilcoxon test $V = 148$, n.s.).

Discussion and Conclusions

- The low d’ scores in the identification task suggest that listeners were unable to categorize /d/-flaps and /t/-flaps.
- While listeners were able to distinguish /d/-tokens from /t/-tokens in the ABX task, they were unable to do this in the AB task where they heard only two sounds per trial, which are never identical.
  - Speakers were unable to use the ‘unrelated cues’ strategy in the AB task, suggesting that listeners cannot distinguish /d/-flaps from /t/-flaps on the basis of cues relevant to the underlying voicing contrast.
  - These results hold in both an identification and a discrimination task, and when frequency effects are mitigated through the use of nonce words.
  - If listeners are neither able to distinguish nor categorize /d/-flaps and /t/-flaps, speakers who maintain this distinction must be doing so for reasons other than listeners’ benefit.

Methods, Part II

- To test whether listeners discriminated based on irrelevant acoustic differences between A and B in the ABX task, an AB Task (A,k.a 2 AFC) was run.
- Allows for comparisons (like the ABX task), but no two tokens are the same on a given trial (like the ID task).
- 21 listeners heard two tokens (from the previous set) per trial.
- Half of the trials had /d/ first and half had /t/ first.

Selected References


Acknowledgements

Thanks are due to Shigeto Kawahara, Bruce Tesar, Kristen Syrett, and the Spring 2010 Rutgers Phonetics Seminar.