Is phonotactic repair of onset clusters modulated by listener expectations?

Max J. Kaplan
University of California, Santa Cruz

UNIVERSITY OF CALIFORNIA SANTA CRUZ

Digital Poster

AMLAP23 *31.8-2.9, 2023*
BCBL, Donastia-San Sebastián, ES

Background.

- Non-native sound sequences often systematically misheard in ways governed by native phonotactics (called ***phonotactic repair***) [2][3][4][5]
  - Mapping of auditory cues affected by task-related expectations (about contrasts present in speech stream) in addition to language experience [2]
- EEG evidence: listeners form early, unconscious abstract representations of illegal sequences which do not arise in linguistic behavioral tasks [1]

➢ **Can unrepaired representations be active in conscious linguistic behavior?**

- Native English & French listeners:

Actual		Reported [5]
*[tl-]	⇨	"KL"
*[dl-]	⇨	"GL"

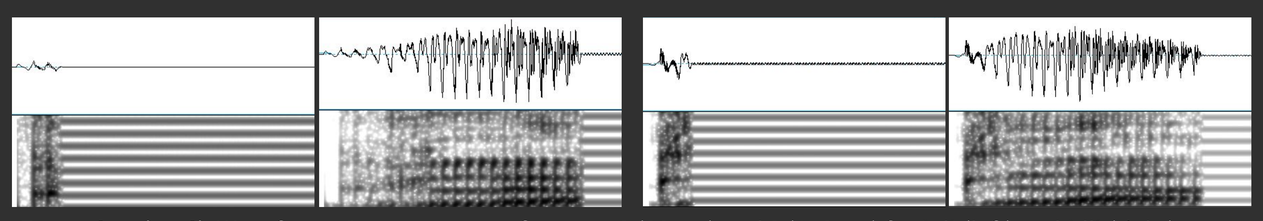
  - Due to spectral similarity of pre-lateral [t] & [k] release bursts [6]
  - French listeners: *less* post-burst information led to greater accuracy in gated transcription task (see *Methods* below): [4]
    - Transcription accuracy only declined after ***vowel*** onset (gate 4, 70 ms)
      Possible effect of syllable structure?
    - Unexpectedly high overall accuracy (~74% at final gate); incongruent with that in ungated forced-choice (~43%) & free transcription (~14%)

➢ **Do these results hold in English? How robust is early gate accuracy?**

Methods.

Materials & Participants

- Native English listeners (N=63), no childhood experience with languages identified as having [tl] or [dl] onsets
- Stimuli: syllables (from [6]), beginning with:
  - Lateral clusters ***[tla]**, ***[dla]**, **[kla]**, **[gla]**; and
  - Glide clusters [twa], [dwa], [kwa], [gwa]
  - **Gated** in 20 ms intervals to include between 10-190 ms following release burst (following [5]; only 30-170 ms gates included in analysis)
  - Final 5 ms ramped to zero; signal replaced with 500 Hz square wave



Example stimuli. Waveforms & spectrograms for gates 2 (30 ms) and 9 (170 ms) for [gl] (left) and [dl] (right).

Procedure

- Online participation with PCIbex [8]
- Three task between-subjects conditions:
  - Free transcription (typed) in text-entry box (following [4])
    - Unguided (N=34): Listeners told stimuli would be "chunks of syllables"
    - Guided (N=14): Listeners told "Some of these syllables will begin with sounds that don't occur together in English."
  - Forced-choice button press (N=15)
    - Mouse-click; all 8 possible onsets presented in writing
    - Listeners told they would hear all 8 sequences
- Practice block with gated [pw] & [bw] stimuli
- 4 randomized test blocks,total of 8 observations per stimulus condition, per participant
- No feedback in any condition

Hypotheses.

Rate of repair

a. Should vary inversely with *plausibility* (≈ baseline activation) & thus likelihood of reaching the activation threshold for selection by the parser
   - Predicts a **cline by condition**, from unguided transcription (most repair) to forced choice (least)

b. Will be **greatest** at **later gates** (after vowel onset)

Results.

- Very low overall accuracy for illicit coronal *[tl-] & *[dl-] stimuli
  - 85.8% of illicit coronal stimuli repaired to velar; just 2.7% for licit coronals
  - No evidence of other significant repairs (e.g., epenthesis to yield [təla])
  - Asymmetric confusability: velar rarely reported as coronal
- Accuracy: somewhat higher in guided tasks (n.s. for transcription: $z = 0.9223$, $p = 0.3564$; significantly for forced-choice: 37:1 odds; $z = 8.6065$, $p < 0.0001$)
- Greatest response accuracy at earlier gates
  - *[tl-] & [dl-] most accurate at gates 2 & 4 in Guided Transcription ($z = 2.1263$, $p = 0.0335$) & Forced-Choice ($z = 3.5828$, $p = 0.0003$)
- Preliminary RT data:
  - Correct "TL"/"DL" responses had greater IQR, while overall RTs for [tl, dl] stimuli were the same as others.

Figure 1. Rates of confusion. Proportion of response onset (x-axis) to each stimulus onset (y-axis), collapsed across all gates, by task condition. High "identity" accuracy for all but the illicit stimuli (outlined in red), which are instead typically misidentified as their velar counterparts. Mean rates of velar-initial response to coronal-initial stimuli were similar in b. Guided Transcription (.846) & c. Forced-Choice tasks (.882) vs. a. Unguided Transcription (.845). Mean identification accuracy was highest in c. (.075), then b. (.029), and lowest in c. (.006).

Figure 2. Rates of repair. Proportion velar-initial responses to illicit *[tl-] & *[dl-] stimuli – by gate & condition. Rates in all conditions were lowest up to the vowel onset (≤ 70 ms). Rates of repair were at floor for licit [tw-] & [dw-] stimuli across all conditions. Y-axis begins at 55%.

Discussion.

- Very high rates of repair & only weak mitigating effect of guidance or task
  - Accuracy greater in forced choice, but similar proportions of repair
- Partially replicates [4]: accuracy greatest at pre-vowel gates
  - Suggests additional auditory information or phonological structure may be detrimental to the activation of illicit representations
  - Higher rate of repair may represent crosslinguistic or methodological differences

Future Directions.

- Crosslinguistic comparison and bilingualism
  - Mexican Spanish contrasts [tl] & [kl]; Mandarin bans both
    - Establishing monolingual baselines (for forced-choice)
    - Testing bilinguals, manipulating language mode (as in [9])
- On-line processing
  - Gating only simulates timecourse of early perception
  - Eyetracking may provide insight about repair timecourse

Selected references. **[1]** Breen, Kingston, & Sanders, 2013. Perceptual representations of phonotactically illegal syllables. **[2]** Brown & Hildum, 1958. Expectancy and the perception of syllables. **[3]** Dupoux, et al., 1999. Epenthetic vowels in Japanese: A perceptual illusion? **[4]** Hallé, et al., 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? **[5]** Hallé & Best, 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. **[6]** Kawasaki-Fukumori, 1992. An acoustical basis for universal phonotactic constraints. **[7]** Massaro & Cohen, 1983. Phonological context in speech perception. **[8]** Simonet & Amengual, 2020. Increased language co-activation leads to enhanced cross-linguistic phonetic convergence. **[9]** Zehr & Schwarz, 2018. PennController for Internet Based Experiments (IBEX).