Listener responses to co-articulation: Solving the "Use it or Lose it" problem

Main session, Phonetics
Anne Pycha[a] and Georgia Zellou[b]
[a]University of Wisconsin, Milwaukee and [b]University of California, Davis

Co-articulation occurs when two articulatory gestures from discrete sounds overlap in time[1]. For example, in English phrases such as *a peep* /ə pip/ vs. *a pope* /ə po^{w}p/, anticipatory vowel-to-vowel co-articulation yields different acoustic realizations of schwa, e.g., [ə$_{i}$] vs. [ə$_{ow}$][2]. Prior work on listener perception of co-articulation has provided evidence for two types of behavior that are potentially in conflict, which we refer to as the "Use it or Lose it" problem. On the one hand, listeners use co-articulatory information to make predictions about upcoming words: upon hearing [ə$_{i}$], for example, they may predict that the next word is *peep*[3–5]. On the other hand, listeners also lose this information: upon hearing [ə$_{i}$ pip] or [ə$_{ow}$ po^{w}p], they may compensate – that is, they may use contextual information to factor out acoustic perturbations, such that [ə$_{i}$] and [ə$_{ow}$] are both perceived as canonical versions of schwa[6–8]. In the current study, we asked if these two behaviors are mutually exclusive. That is, if a listener loses co-articulatory information through compensation, does this mean he or she cannot use it to make predictions?

We addressed this question by asking listeners to respond to the same set of acoustic tokens in two tasks: one that quantified prediction (task 1), and another that quantified compensation (task 2). Test stimuli consisted of five phrases, *a peep, a pope, a poop, a pop, a pipe,* in which the schwa exhibited effects of vowel-to-vowel co-articulation ([ə$_{i}$], [ə$_{ow}$], [ə$_{u}$], [ə$_{ɑ}$], [ə$_{ai}$]; same-spliced from different productions of each utterance). Control stimuli contained non-coarticulated schwa (cross-spliced from *a pup*). Task 1 used a visual-world paradigm with eye-tracking, as in Figure 1.

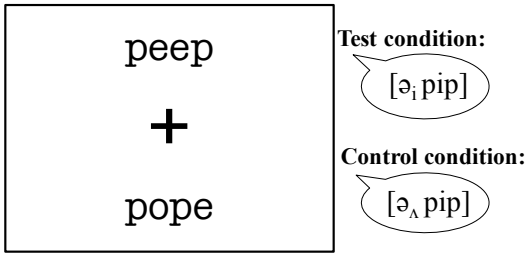

FIGURE 1. TASK 1, PREDICTION.

Test condition:
[ə$_{ʌ}$pʌp ə$_{i}$pip] vs. [ə$_{ʌ}$pʌp ə$_{ʌ}$pip]

Control condition:
[ə$_{i}$pip ə$_{ʌ}$pip] vs [ə$_{i}$pip ə$_{i}$pip]

Task: Which pair, first or second, contains schwas that are most different from each other?

FIGURE 2. TASK 2, DISCRIMINATION

On each trial, listeners first visually previewed orthographic target and competitor words. Subsequently, they were instructed to look at a fixation cross, then to listen to a spoken stimulus (e.g., *now look at* [ə$_{i}$ pip]). For test trials, faster looks to target words would indicate that co-articulation facilitates word perception (using). Task 2 was a paired discrimination task, as in Figure 2. On each trial, listeners heard two pairs of phrases and indicated which pair contained schwas that were most different from one another (design from [3]). One pair always contained acoustically identical (non-coarticulated) schwas and the other contained different schwas (coarticulated vs. non-coarticulated). Test trials consisted of pairs with different words (e.g., [ə$_{ʌ}$pʌp ə$_{i}$pip] vs. [ə$_{ʌ}$pʌp ə$_{ʌ}$pip]); control trial pairs contained the same words (e.g., [ə$_{ʌ}$pip ə$_{i}$pip] vs. [ə$_{i}$pip ə$_{i}$pip]). Compared to control trials, good performance on test trials would indicate that perception is veridical. However, poor performance would indicate that compensation occurs (losing), i.e., that all four realizations of schwa are perceived as canonical. Crucially, we hypothesized that listeners' performance across these two tasks must be logically related: if listeners tend to lose co-articulatory information through compensation, then they cannot also use this information for prediction.

Preliminary results from ten participants provide support for this hypothesis. A mixed-effects logistic regression model revealed that looking times (averaged by-word and by-subject) in task 1 significantly predicted discrimination accuracy in task 2 (β = -0.52, std. error = 0.21, z = -2.50, p = 0.01). The correlation between looking time and discrimination was negative, as seen in Figure 3. That is, when listeners took longer to use the coarticulatory information on schwa to direct their eye fixation to a target word, their accuracy in discriminating between coarticulated versus non-coarticulated schwas tended to be lower. This suggests that compensation leads to the loss of co-articulatory information that could otherwise be used for prediction. This finding connects two previously disparate strands of work on the perception of coarticulation, and, as far as we know, it is the first to address the "Use it or Lose it" puzzle.

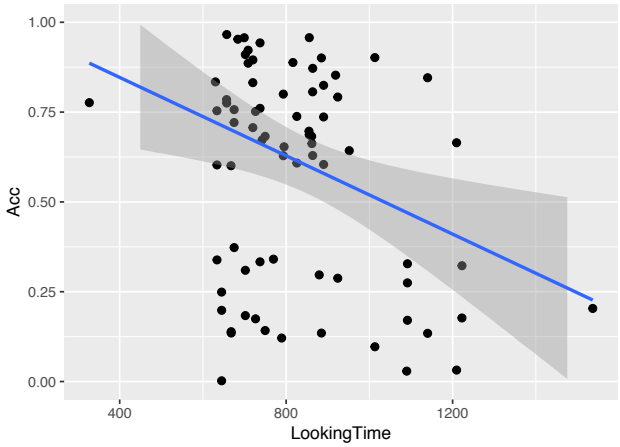

FIGURE 3. RESULTS. SIGNIFICANT NEGATIVE CORRELATION BETWEEN ACCURACY IN DISCRIMINATION TASK (Y-AXIS, JITTERED) AND LOOKING TIMES TO VISUAL TARGETS (IN MS, X-AXIS).

1. Byrd, D. Perception of Assimilation in Consonants Clusters:A Gestural Model. *PHO* **49,** 1–24 (1992).
2. Magen, H. S. The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics* **25,** 187–205 (1997).
3. Beddor, P. S., Harnsberger, J. D. & Lindemann, S. Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics* **30,** 591–627 (2002).
4. Beddor, P. S., McGowan, K. B., Boland, J. E., Coetzee, A. W. & Brasher, A. The time course of perception of coarticulation. *The Journal of the Acoustical Society of America* **133,** 2350–2366 (2013).
5. Dahan, D., Magnuson, J. S., Tanenhaus, M. K. & Hogan, E. M. Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes* **16,** 507–534 (2001).
6. Beddor, P. S. & Krakow, R. A. Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *The Journal of the Acoustical Society of America* **106,** 2868–2887 (1999).
7. Ohala, J. J. The phonetics of sound change. in *Historical linguistics: Problems and perspectives* 237–278 (1993).
8. Yu, A. C. L. & Lee, H. The stability of perceptual compensation for coarticulation within and across individuals: A cross-validation study. *The Journal of the Acoustical Society of America* **136,** 382–388 (2014).