Dynamic hyper-articulation: Effects of context and lexical representations
Esteban Buz1*, Scott Seyfarth2, T. Florian Jaeger1
1University of Rochester, 2University of California, San Diego
* corresponding author: ebuz@ur.rochester.edu

Speakers hyper-articulate words when phonetically-similar alternatives are salient in the speech context. For example, Baese-Berk and Goldrick (2009) find that speakers produce the [pʰ] in *pill* with a longer voice-onset time (VOT) when its minimal pair (MP) *bill* is contextually co-present (see also Buz, Jaeger, & Tanenhaus, 2014; Kirov & Wilson, 2012). The underlying source of this hyper-articulation is an area of active debate. Some proposals attribute hyper-articulation to competition between lexical and phonological representations during linguistic encoding (Baese-Berk & Goldrick, 2009; Fox, Reilly, & Blumstein, 2015; Fricke, Baese-Berk, & Goldrick, 2016; Goldrick, Vaughn, & Murphy, 2013; Kirov & Wilson, 2012). Other proposals attribute this hyper-articulation to speakers selectively enhancing phonetic contrasts to achieve communicative goals (cf. Jaeger, 2013; Lindblom, 1990; Schertz, 2013). While both accounts predict phonetic effects on contrastive segments, they differ in two important respects. First, competition-based accounts predict consistent change in speech regardless of the type of contrast (Baese-Berk & Goldrick, 2009; Goldrick et al., 2013), whereas enhancement-based accounts predict an effect that varies based on what will exaggerate the perceptual distance between the target word and its alternative. Second, competition-based exaggeration is lexically-mediated: regardless of the context, it predicts hyper-articulation of words with MPs relative to phonetically-matched words without MPs, although hyper-articulation may increase further when the MP is relevant in the context (Baese-Berk & Goldrick, 2009). On the other hand, most enhancement-based accounts are context-dependent: hyper-articulation only occurs when an MP is relevant (Lindblom, 1990). We test these predictions by investigating whether hyper-articulation is sensitive to perceptual distance or is primarily determined by segmental position (Study 1), and whether lexical factors influence articulation (Study 2).

**Study 1, positional vs. perceptual exaggeration** To evaluate whether hyper-articulation is contingent on the perceptual cues associated with the phonetic contrast, or on the position of the contrast within a word, we recruited 40 participants to play an online communication game with a (simulated) partner. In each trial of the game, the participant saw three words on their computer screen. One word was highlighted, and the participant read the highlighted target word aloud in order to direct their partner to select it from among the three alternatives.

The 36 critical target words comprised 18 English MPs that differed on the voicing of a word-final sibilant /s, z/ (e.g., *dose* and *doze*). On CONTRASTIVE trials, one of the alternative words that appeared next to the target was its MP: e.g., if *dose* was highlighted, one of the alternatives was *doze*. On CONTROL trials, both of the alternative words were frequency-matched fillers. This design has previously been shown to elicit hyper-articulated productions of the target word in contrastive trials, relative to control trials (Baese-Berk & Goldrick, 2009, et seq.).

In American English, the major cues that distinguish coda [s] and [z] are the durations of the nucleus vowel and of the sibilant itself: vowels are shorter before [s] than before [z], and [s] is longer than [z] (Klatt, 1976; Raphael, 1970). Therefore, enhancement-based accounts predict different phonetic effects for the production of *dose* versus *doze*: speakers should shorten vowels before [s] and lengthen [s] in *dose*, while lengthening vowels before [z] and shortening [z] in *doze*. Alternatively, the competition-based account predicts that the effects on [s]-final *dose* and [z]-final *doze* should be the same: there should be either elongation of segments in both *dose* and *doze* (Baese-Berk & Goldrick, 2009), or shortening in both kinds of words (Goldrick et al., 2013) but not differential elongation and shortening.

We found that vowel durations were significantly shorter in the contrastive condition relative to the control condition in [s]-final words ($\hat{\beta} = -9\text{ms}; p < 0.05$, by maximal LMEM), but vowel durations in [z]-final words did not significantly differ ($\hat{\beta} = +1\text{ms}; p > 0.05$). We found no effect of condition on sib-
lant durations. In a follow-up analysis, we calculated the duration of voicing during each sibilant (Stevens, Blumstein, Glicksman, Burton, & Kurowski, 1992). We found that voicing during [z] was extended significantly longer in the contrastive condition ($\hat{\beta} = +16 ms; p < 0.001$), but that voicing during [s] did not significantly differ.

Our results support the hypothesis that hyper-articulation is based on the perceptual contrast, rather than on lexical competition. The differential effects (shortening of vowels and increased voicing) both enhance the contrast between the target word and its MP alternative. This is contra the predictions of consistent changes (i.e. either elongation or shortening), that is predicted by competition-based accounts (cf. Goldrick et al., 2013).

**Study 2, Lexical factors** Under competition interpretations of hyper-articulation, the underlying source of increased competition in the co-activation of alternative lexical and phonological representations. Under some accounts, competitors are always co-active and can influence linguistic encoding even when not co-present. In favor of this is evidence that having an MP in the lexicon (e.g. *bill*) affects the articulation of a target word (e.g. *pill*), even when it is not present, relative to similar words without an MP (e.g. *pipe*, see Baese-Berk & Goldrick, 2009; Fox et al., 2015). Nevertheless, evidence that word-final MPs affect articulation is less clear (Goldrick et al., 2013). Study 1 found that speakers hyper-articulate word-final contrasts when a voicing contrastive MP is contextually co-present; thus, a critical follow-up question for the competition-based account is whether words with voicing MPs (e.g. *bays*) are hyper-articulated relative to words without voicing MPs (e.g. *daze*).

To evaluate whether articulation is mediated by lexical factors in the absence of a contextually relevant MP, we recruited 60 participants for the same communication game used in Study 1. The critical targets were pairs of matched sibilant-final words that either did or did not have a voicing-final MP (e.g., *bays* has MP *base*, but phonologically-matched *daze* has no MP *dase*). Target pairs were matched on segment length, frequency and phonological differences that would bias vowel or sibilant durations. Of the 18 critical pairs, half were voiced (*bays–daze*) and half were voiceless (*pace–case*). Participants produced both members of each pair during the course of the experiment, allowing for both within-participant and within-item comparison. We found no differences in vowel, sibilant, or voicing durations across targets with and without MPs ($p's > 0.1$).

**Conclusions** We find that hyper-articulation is consistent with enhancement-based accounts that predict effects that would exaggerate the perceptual distance between a target and its competitor, and not with competition-based accounts that predict perception-insensitive effects on production. Further, we find no evidence that lexical factors are the source of this hyper-articulation (contra Baese-Berk & Goldrick, 2009; Goldrick et al., 2013). The contrast-specific and contextual-dependent nature of these changes suggest that hyper-articulation may be communicative, rather than derived from speaker-internal lexical competition.