ELSEVIER
Cognitive Science 27 (2003) 801–805
COGNITIVE SCIENCE
http://www.elsevier.com/locate/cogsci

Letter to the Editor

Lexical effects on compensation for coarticulation: a tale of two systems?

James S. Magnuson[a,*], Bob McMurray[b],
Michael K. Tanenhaus[b], Richard N. Aslin[b]

[a]*Department of Psychology, Columbia University, 1190 Amsterdam Ave., MC5501, New York, NY 10025, USA*
[b]*Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 10025, USA*

Received 13 May 2003; received in revised form 3 July 2003; accepted 8 July 2003

Abstract

We reply to McQueen's commentary by comparing the parsimony of his account of relevant data and the computational model he favors with the explanation and model we favor. His account requires multiple independent explanations and mechanisms. Ours requires one: lexical feedback.

Keywords: Lexical effects; Parsimony; Computational model; Language understanding

McQueen's account of lexically-mediated compensation-for-coarticulation (LCfC), in his commentary on our article (Magnuson, McMurray, Tanenhaus, & Aslin, 2003; MMTA hereafter), requires that data from studies demonstrating LCfC be divided into three sets, each requiring an independent explanation: (1) all other positive evidence of LCfC aside from ours depends upon unspecified high order transitional probabilities (TPs), (2) MMTA's results are attributed to perceptual learning, and (3) the absence of LCfC despite higher-order TP biases in Pitt and McQueen (1998; PM hereafter) requires an as yet unspecified additional explanation, as it contradicts the higher-order TP explanation proposed for other results. We will review each of these briefly, introducing new analyses that challenge McQueen's higher-order TP and perceptual learning explanations, and then conclude with a comparison of the multiple computational mechanisms McQueen proposes to account for the data and the single mechanism required of our account.

---

[*] Corresponding author. Tel.: +1-212-854-5667; fax: +1-212-854-3609.
E-mail address: magnuson@psych.columbia.edu (J.S. Magnuson).

0364-0213/$ – see front matter
doi:10.1016/S0364-0213(03)00067-3

1. Three data sets, three explanations

1.1. Explanation 1: all positive LCfC results except those in MMTA

McQueen speculates that "higher-order TPs could underlie all other apparent lexical effects in compensation for coarticulation" aside from those reported in MMTA. Higher-order TPs are needed because MMTA demonstrated in corpus analyses that claims that Elman and McClelland (1988) confounded lexical and diphone biases (e.g., Cairns, Shillcock, Chater, & Levy, 1995) are incorrect. McQueen did not provide analyses to support the higher-order TP hypothesis, so we conducted a further corpus analysis to examine whether a context of a fixed size could account for the LCfC effects in the literature. We determined the crucial *n*-phone (where *n* is the number of preceding phones + target phone) for 13 of the 16 words for which LCfC has been found[1] (we excluded MMTA's items, and *fish*, which, like *brush*, violates the hypothesis; it is s-biased at the diphone and equi-biased at the triphone level). The value of *n* ranged from 2 to 5 ($M = 3.1$, $SD = 1.22$). To normalize for word length (ranging from 3 to 10 phonemes), we recomputed *n* as the proportion of word length required to make the correct prediction. The range was 0.2–0.83 ($M = 0.46$, $SD = 0.24$). Given the variability of the *n* required for different items, there is no plausible single *n*, as a raw value or proportion, to support the higher-order TP hypothesis.[2] We have not exhausted the possible relevant contexts, but this analysis suggests a higher-order account will be difficult to construct, if not untenable. We suspect that word-specific *n* is governed by a complex relationship among neighborhoods of words, which is easily conceptualized and simulated via lexical feedback in TRACE.

1.2. Explanation 2: absence of LCfC inPitt and McQueen (1998)

Higher-order TPs are invoked to account for "all other apparent [LCfC] effects," but they cannot account for the absence of LCfC for PM's contexts *jui*(*ce*) and *bu*(*sh*). While PM reported these were equi-biased at the diphone level, they are not at the triphone level: /ĵu/ is s-biased (p(s|ĵu) $= .0283$, p(ʃ|ĵu) $< .0001$), and /bʊ/ is ʃ-biased (p(s|bʊ) $< .0001$, p(ʃ|bʊ) $= .0699$). McQueen has not explained why diphone TPs dominate for PM's contexts, while higher-order TPs dominate for "all other apparent [LCfC] effects," but he is concerned that our account does not provide an explanation for the absence of LCfC with PM's lexical contexts.[3] In fact, we cited them as our motivation and explained that we consider them anomalous, based on the results of Samuel and Pitt (2003). They examined perceptual grouping factors that modulate CfC, including the degree to which a fricative is bound to the preceding syllable (Mann & Repp, 1981). The more closely bound the fricative, the weaker the CfC effects. Samuel and Pitt demonstrate that PM's lexical items bind the fricative more strongly than their non-words, thus reducing the opportunity for CfC. The absence of lexical effects in PM likely resulted from a failure to establish the conditions necessary for the more basic CfC effect.

1.3. Explanation 3: MMTA's data

MMTA demonstrated LCfC using stimuli for which there was no possible TP explanation.[4] McQueen suggests that perceptual learning might explain the results, without invoking feedback during processing. He argues that subjects could have learned between-phoneme dependencies as the experiment progressed because the stimuli did not include non-word endpoints. A logical prediction from this explanation is that lexical effects should emerge and strengthen throughout the experiment, as subjects gain enough experience with the materials to detect the missing endpoints. In fact, the magnitude of the lexical effect did not differ in the first and second halves of the experiment. The simple effect of lexical context was of similar size in each half (first: $F(1, 15) = 4.69$, $p = .047$, $\omega^2 = 0.103$; second: $F(1, 15) = 3.81$, $p = .070$, $\omega^2 = 0.081$), and there was not a significant interaction of half with lexical context ($F(1, 15) = 0.13$, $p = .720$, $\omega^2 = 0$). In a more fine-grained analysis, in which we divided the experiment into 6 epochs of 54 trials (all stimuli were repeated every 54-trials, in random order), the interaction of lexical context and epoch was still not reliable ($F(5, 75) = 1.57$, $p = .18$, $\omega^2 = 0.015$). The simple effect of lexical context was not reliable in any 1 epoch, but differences in "*k*"-response rate by lexical context at each epoch were 0.03, 0.06, 0.09, 0.01, 0.13, and 0.03, respectively. The fourth and sixth epochs show that the small difference in epoch 1 cannot be attributed to lack of experience with the experimental materials.

An alternative is that perceptual learning was too fast to be detected. However, for the perceptual learning hypothesis to be viable, between-phoneme interactions must *stably* capture as much of the variance as possible. A system that changes too rapidly risks over-fitting the data and losing useful generalizations for which there has been no recent evidence. Our analyses cannot rule out extremely rapid perceptual learning, of the kind that would be problematic for just this reason. However, they *would* have been sensitive enough to detect more useful perceptual learning on the timescale that has been observed in experiments showing implicit learning from brief exposure to auditory inputs (e.g., Onishi, Chambers, & Fisher, 2002).

2. The tale of two systems

A new theoretical position about feedback (Norris, McQueen, & Cutler, in press) is central to McQueen's commentary. Norris et al. now acknowledge a helpful sublexical role for lexical knowledge, but propose a distinction between on-line feedback and off-line, "for learning" feedback, in analogy to the distinction in neural networks between feedback during activation and weight changes via backpropagation. Norris et al. propose adding a "for-learning" mechanism to their feed-forward model, Merge, to account for lexical effects (e.g., LCfC) that cannot be handled by its post-lexical phonemic decision nodes.[5] Thus, in order to avoid having feedback directly modulate sublexical processing, two separate mechanisms are needed. In contrast, on-line feedback provides an implicit encoding of the context-specific prior probabilities afforded by Merge's proposed prelexical mechanism, while simultaneously accounting for the effects motivating Merge's post-lexical phoneme nodes.

To conclude, McQueen's feedback for perceptual learning account of the MMTA data is not supported by our new analyses. If extremely rapid perceptual learning is the correct account, then a study explicitly designed to reveal it is required. McQueen's speculation that higher-order TPs would account for related results is vague, it is contradicted by PM's failure to find LCfC using two lexical contexts with higher-order TP biases, and our new corpus analyses suggest such an account will be difficult to construct, if not untenable. Parsimony favors the feedback account, both in terms of providing a coherent account of the extant data, and the lesser complexity it requires of a computational model.

Notes

1. There have been positive results with 16 contexts (6 [Elman & McClelland, 1988] + 8 [Samuel & Pitt] + 2 MMTA), and failures with 8 (2 [PM] + 6 [Samuel & Pitt, 2003, who also provide acoustic/perceptual explanations for all 8 failures]).
2. One could not set n to be the maximum observed (e.g., 5), as this would not account for words shorter than n phonemes.
3. McQueen's characterization of PM's results as a dissociation of lexical bias and TP seems based on an expectation that feedback must predict CfC when there is lexical bias in fricative labeling. The logic of LCfC does not force us to consider feedback as the only explanation for Ganong (1980) effects in fricative labeling. The purpose of the paradigm is to distinguish perceptual from post-perceptual effects. Indeed, the flat fricative response rates we (and others) find across *t*/*k* steps for particular fricatives as "*k*"-response rate varies shows the two are not yoked.
4. An aside about our materials: McQueen questions whether our stimulus *brush* has the same vowel as *christmas*; for our speaker in citation context, it has. He also worries that /Λ/ and /ə/ are not distinguished in our corpora; they are, and both are s-biased: p(ʃ|ə) = .0014, p(s|ə) = .0211; p(ʃ|Λ) =.044; p(s|Λ) = .0892. Whether or not *brush* is the ideal foil for *christmas*, it is ideal for pitting lexical context against TP.
5. McQueen also cites previous simulations of LCfC with SRNs without explicit lexical representations as evidence that on-line feedback is not required to account for LCfC. We think this argument stems from a misunderstanding of information flow in SRNs (see MMTA, note 7), which can be clarified by the distinction between on-line and for-learning feedback: SRNs have both. There is hidden-to-hidden unit on-line feedback and "for-learning" weight change via backpropagation. Thus, the simulations do not show that on-line feedback is not required, only that another architecture employing on-line feedback can simulate LCfC.

Acknowledgments

The preparation of this manuscript was supported by NIDCD DC-005765 to JSM, and NIDCD DC-005071 to MKT and RNA. We thank Gary Dell and James McQueen for helpful comments.

References

Cairns, P., Shillcock, R., Chater, N., & Levy, J. P. (1995). Bottom-up connectionist modeling of speech. In J. P. Levy & D. Bairaktaris (Eds.), *Connectionist models of memory and language* (pp. 289–310). London, England: UCL Press Limited.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, *27*, 143–165.

Ganong, W. F. (1980). Phonetic categorization in auditory perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 110–125.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science*, *27*, 285–298.

Mann, V. A., & Repp, B. H. (1981). Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America*, *69*, 548–558.

Norris, D., McQueen, J. M., & Cutler, A. (in press). Perceptual learning in speech. *Cognitive Psychology*.

Onishi, K. H., Chambers, K. E., & Fisher, C. L. (2002). Learning phonotactic constraints from brief auditory exposure. *Cognition*, *83*, B13–B23.

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, *39*, 347–370.

Samuel, A. G., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, *48*, 416–434.