COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 47 (2023) e13342

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13342

Lexically Mediated Compensation for Coarticulation Still as Elusive as a White Christmash

James M. McQueen,[a] Alexandra Jesse,[b] Holger Mitterer[c]

[a]*Donders Institute, Radboud University*
[b]*Department of Psychological and Brain Sciences, University of Massachusetts*
[c]*Department of Cognitive Science, University of Malta*

Received 4 November 2022; received in revised form 20 July 2023; accepted 30 August 2023

Abstract

Luthra, Peraza-Santiago, Beeson, Saltzman, Crinnion, and Magnuson (2021) present data from the lexically mediated compensation for coarticulation paradigm that they claim provides conclusive evidence in favor of top-down processing in speech perception. We argue here that this evidence does not support that conclusion. The findings are open to alternative explanations, and we give data in support of one of them (that there is an acoustic confound in the materials). Lexically mediated compensation for coarticulation thus remains elusive, while prior data from the paradigm instead challenge the idea that there is top-down processing in online speech recognition.

Keywords: Cognitive penetrability; Language; Speech perception; Lexical effects; Feedback

1. Introduction

An old but fundamental question in cognitive science is whether knowledge held at a later stage of perceptual processing influences, through top-down and online feedback, mental operations taking place at an earlier processing stage. This question is fundamental because the answer to it places important constraints on the nature of cognitive processing: are there or are there not mental modules that are cognitively impenetrable (Fodor, 1983)?

Correspondence should be sent to James M. McQueen, Donders Institute, Radboud University, Thomas van Aquinostraat 4, 6525 GD Nijmegen, The Netherlands. E-mail: james.mcqueen@donders.ru.nl

In the domain of speech perception, the search for such online feedback effects has focused on the lexically mediated compensation for coarticulation (LCfC) paradigm. The focus on this paradigm has been because, as Elman and McClelland (1988) argued, the paradigm could provide convincing evidence of top-down processing. Specifically, it has the potential to show that higher-level knowledge about the phonological content of words can influence lower-level compensation processes that modulate the interpretation of acoustic information. Elman and McClelland claimed to present such evidence. Since then, however, a variety of alternative explanations for the original findings have been proposed, controlled for, and/or tested directly. The debate has thus gone back and forth, with some LCfC studies making the case for and some against top-down processing.

Luthra et al. (2021) is the latest round in this debate. Luthra et al. conclude in favor of top-down processing. We argue here that, in spite of the many positive features of their study, including preregistration, replication, open data, and extensive stimulus piloting, their conclusion is not warranted, for four reasons.

2. Four counter-arguments

2.1. Evidence against top-down processing

Evidence from the LCfC paradigm that lexical effects dissociate (McQueen, Jesse, & Norris, 2009; Pitt & McQueen, 1998) is inconsistent with top-down processing. Although Luthra et al. (2021) cite these studies, they do not discuss these dissociations. In the studies, listeners made four-alternative forced-choice (4AFC) decisions about the final fricatives at the ends of the first words in the two-word stimulus sequences and the two initial stops in the second words (i.e., [st], [sk], [ʃt], or [ʃk] in "*Christma*[s/ʃ] [t/k]*apes*"). Listeners used their lexical knowledge to make decisions about the ambiguous fricatives (i.e., there was a Ganong effect), while on the same trials (i.e., in the same 4AFC responses) they did not use this knowledge to modulate their interpretation of the following stops (i.e., there was no LCfC effect). In other trials (intermixed in the running order of the experiments), however, there were CfC effects with stimuli from the same fricative and stop continua. The two components of the LCfC effect (a Ganong effect and a CfC effect) were thus present in the experiment but on the critical trials, they dissociated.

This dissociation challenges top-down processing because, according to the theory, both effects are caused by lexical-prelexical feedback. That is, ambiguous fricatives that are sufficiently lexically biased to produce a Ganong effect should also cause LCfC, at least if the experimental conditions (e.g., stimuli, task) are otherwise right for a CfC effect. Those conditions are met if, within a study showing the dissociation (the word-context conditions of Pitt & McQueen, 1998; Experiment 3 in McQueen et al., 2009), there are other parts of the study where CfC is caused by ambiguous fricatives biased in nonlexical ways (e.g., by transitional probability [TP] biases in the nonword-context conditions of Pitt & McQueen, 1998; by experiment-induced biases in Experiment 1 in McQueen et al., 2009). In a study in which those conditions are met, top-down processing predicts that Ganong and LCfC effects should co-occur.

Samuel and Pitt (2003) argued that the dissociation in Pitt and McQueen (1998) could be the result of effects of perceptual grouping. McQueen et al. (2009), therefore, controlled for effects of perceptual grouping (see Supplementary Materials for further discussion), and yet the dissociation was found in three experiments.

One might argue instead that these dissociations are based, in part, on null results: the failure to observe LCfC effects. Perhaps such effects were present in these situations but they are too weak to be detected. To address this issue, we carried out Bayes Factor (BF) analyses of the data in Experiment 3 in McQueen et al. (2009). This experiment was a replication of Magnuson, McMurray, Tanenhaus, and Aslin (2003), but the experimental bias that had been introduced during the practice trials of the original study was removed. A Ganong effect was observed, but no LCfC effect with the same materials, in the same 4AFC responses. These BF analyses allowed us to test the strength of the evidence in the data not only for the hypothesis that there is top-down processing but also for the null hypothesis that there is not. The logic of these analyses is that of the LCfC paradigm itself and derives from the theory that there is lexical feedback. Because the Ganong effect and the LCfC effect are assumed to have the same common cause, namely, lexical feedback, the size of the Ganong effect should predict the size of the LCfC effect. According to the theory, as the amount of feedback increases, the Ganong effect should become larger and so should the LCfC effect (as also argued by Samuel & Pitt, 2003, p. 429). Similarly, if there is no Ganong effect, there is no reason to expect an LCfC effect; as Luthra et al. (2021) put it, "we set out to test LCfC only after establishing that we could detect Ganong effects with candidate context items" (p. 4).

Four BF analyses are reported in detail in the Supplementary Materials: three individual ones on the separate ambiguous fricative conditions in the experiment and a combined one. Two of the individual analyses are inconclusive but one shows that the null hypothesis is more than three times more likely than the feedback hypothesis. The combined analysis, which carries the most weight because it is based on more data, shows that the null hypothesis is eight times more likely. This is substantial evidence for the null (Dienes, 2014).

There is, to date, no explanation offered for how these dissociations are compatible with online, top-down feedback in speech perception. The new BF analyses further strengthen the case that such dissociations are evidence against feedback. Before Luthra et al. (2021) can conclude in favor of top-down processing based on their data from the LCfC paradigm, therefore, they need to offer a convincing explanation for other data from the paradigm that contradict their account.

2.2. *Transitional probabilities*

Luthra et al. (2021) consider the possibility (raised, e.g., by Pitt & McQueen, 1998) that apparent LCfC effects may reflect not lexical knowledge per se but rather TPs between segments in the critical words (e.g., the probability of [s] after schwa in *Christmas*). They do not address this issue, however, by controlling for TPs in their stimuli. Indeed, in three of their four conditions, there is a diphone TP bias in the same direction as the lexical bias (for *isolate*, *maniac*, and *questionnaire*, but not for *pocketful*; see the Appendix in Luthra et al.). Furthermore, in three conditions (all except *questionnaire*), there is a triphone TP bias in the same direction as the lexical bias and there are higher-order TP biases in the same direction as the lexical bias in all four conditions. Future research is required to establish what kinds of TP biases (i.e., diphone vs. higher-order biases) may modulate CfC and hence whether LCfC can be detected when TPs are fully controlled. Given the correlation between TP and lexical biases and other issues concerning the computation of TPs (see McQueen et al., 2009), however, this line of research is difficult to pursue. Nevertheless, there is prior evidence that TP biases can potentially underlie apparent LCfC effects (Pitt & McQueen, 1998).

Instead of controlling for TPs, Luthra et al. (2021) argue that TP biases could predict fewer than half of the apparent positive LCfC effects in the published literature and hence that lexical bias is more plausibly the source of those effects. The problem here is that, while this analysis suggests that TP is unlikely to be the sole predictor of the effects across all existing studies, it does not show that TP plays no role in any apparent LCfC effect in any given study, especially given the (replicated) evidence that TP biases, in the absence of lexical biases, can modulate CfC after ambiguous fricatives (Pitt & McQueen, 1998). Furthermore, the apparent positive LCfC effects in the studies where TP is a poor predictor could have arisen for reasons other than lexical bias (e.g., learning based on experiment-internal biases; McQueen, 2003; McQueen et al., 2009). Luthra et al.'s analysis showing that TP is a poor predictor in some studies, therefore, does not show that the apparent LCfC effects in those studies must be due to lexical biases. Their analysis also does not show that apparent LCfC effects in studies with TP biases cannot be due to those biases. This means that the apparent LCfC effects observed by Luthra et al. could be due to the TP biases in their materials.

2.3. *Experiment-induced bias*

Luthra et al. (2021) attempt to address the concern that apparent LCfC effects may be due to offline perceptual learning (McQueen, 2003) rather than online top-down processing. Based on experiment-internal biases, participants in Magnuson et al. (2003) could learn that some interpretations of the ambiguous sounds at the end of the first word in the two-word sequences were more likely than others. During the main experiment, participants heard, for example, ambiguous *Christma*[s/ʃ] and unambiguous *Christmas* but not unambiguous *Christmash* (though note these are not the actual stimuli). During the practice block, they heard, for example, only unambiguous *Christmas*. They thus could learn that the ambiguous sound was more likely to be lexically consistent [s]. This is an example of the phenomenon that the relative proportions of different types of trial in phonetic categorization experiments influence performance (Bushong & Jaeger, 2019; Repp & Liberman, 1987). McQueen et al. (2009) presented evidence to suggest that the participants in Magnuson et al. (2003) did indeed learn from the bias in the practice blocks: with the same stimuli and new participants, the effects reversed when the bias was reversed and went away when the bias was removed. Strikingly, these different outcomes arose on the basis of changes in the content of only 16 trials. Different biases arose depending on whether the practice trials had only words (eight each of, e.g., *Christmas* and *foolish*) or only nonwords (eight each of *Christmash* and *foolis*).

Luthra et al. (2021) attempted to deal with this issue by using only context words with ambiguous final sounds (e.g., *Christma*[s/ʃ], though again this is not actually one of their words). But this does not solve the problem. Given this exposure, and especially that there are no unambiguous nonwords, the participant can still learn that the ambiguous sounds in the context words are the lexically consistent ones. This situation is similar to that in the lexically guided perceptual learning paradigm (Norris, McQueen, & Cutler, 2003), in which exposure to an ambiguous sound in only 10 lexically biased contexts (Kraljic & Samuel, 2007; Poellmann, McQueen, & Mitterer, 2011) is enough for participants to learn that the sound is the lexically consistent one. What is required to prevent this kind of learning is exposure to unambiguous tokens of both interpretations of the ambiguous sound (e.g., *Christma*[s/ʃ] with *Christmas* and *Christmash*, in equal proportions; *fooli*[s/ʃ] with *foolis* and *foolish*). Without control over the proportions of these different trial types, it is impossible to say whether effects are the result of online lexical feedback or of experiment-internal learning. This means that the apparent LCfC effects observed by Luthra et al. could after all be due to experiment-induced bias.

We tested this hypothesis by reanalyzing the data from the two experiments in Luthra et al. (2021), asking whether the apparent LCfC effects changed over the course of the experiment. No interaction of the apparent LCfC effect with trial order was observed in either dataset (see Supplementary Materials). These null results, however, are inconclusive. It could be the case that there was no experiment-internal learning, but it could also be the case that the learning took place so early in the experiment that it could not be detected. Although Luthra et al. had no practice trials, learning could have arisen during the first few trials of the first 100-trial block. Rapid learning is quite likely given that the bias effects in McQueen et al. (2009) arose on the basis of only 16 practice trials and that lexically guided perceptual learning can arise after only 10 trials (Kraljic & Samuel, 2007; Poellmann et al., 2011).

2.4. *Possible acoustic confounds*

The ambiguous versions of Luthra et al.'s (2021) context stimuli (i.e., the first words in the two-word sequences) were made separately (e.g., *isola*[t/k] was made from recordings of *isolate* and the matched nonword *isolake*, while *mania*[t/k] was made from *maniac* and *maniat*). The final sounds were, therefore, not physically identical across the pairs of context words. In the pair based on the endpoints *pocketful* and *questionnaire*, *pocketful* is unvoiced for the last 40 ms, while *questionnaire* is not (see Fig. 1 and the Supplementary Materials for a suggested reason for this difference in voicing). In the other pair (based on the endpoints *isolate* and *maniac*), the stop in the *isola-* context is more like lexically consistent /t/ (it has more energy above 4 kHz; see Fig. 2A), while the stop in the *mania-* context is more like lexically consistent /k/ (it has a maximum at 2.4 kHz, as is typical for a velar pinch). This means that the apparent LCfC effects observed by Luthra et al. could instead be the result of these acoustic differences.

The stimulus construction procedure used by Luthra et al. (2021), therefore, was not adequate. The continuum for each word (e.g., *isolate-isolake*) was made separately, and then an ambiguous continuum step was selected for which there was a lexical bias in the first pilot experiment, for example, more /t/ responses to that step in the *isolate-isolake* (word-nonword)

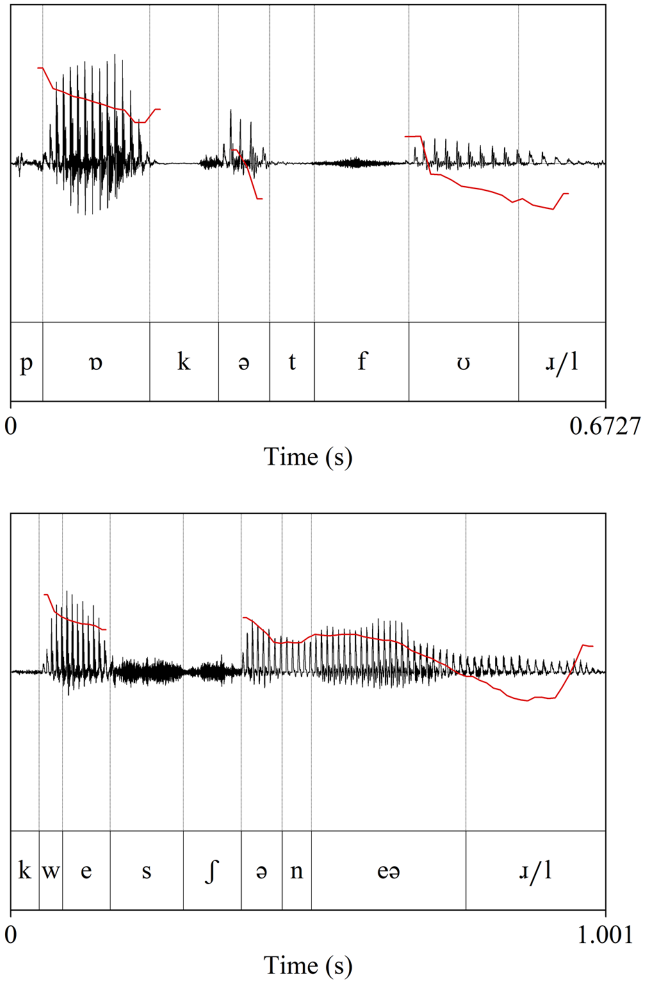

Fig 1. Oscillograms with overlaid pitch contour of the two context stimuli *pocketful* and *questionnaire*, showing devoicing in the *pocketful* ambiguous stimulus and not in the *questionnaire* stimulus.

context than in a trimmed *ate-ake* (nonword-nonword) context. While this method is important in establishing a Ganong effect for the ambiguous context words, because the Ganong effect is a prerequisite for LCfC, it does not guarantee that there are no acoustic differences in the final consonants across the pairs of context words.

3. Cross-splicing experiment

We now ask whether these acoustic differences could underlie the apparent LCfC effects in Luthra et al. (2021). The Luthra et al. pretest data with the ambiguous context sounds in

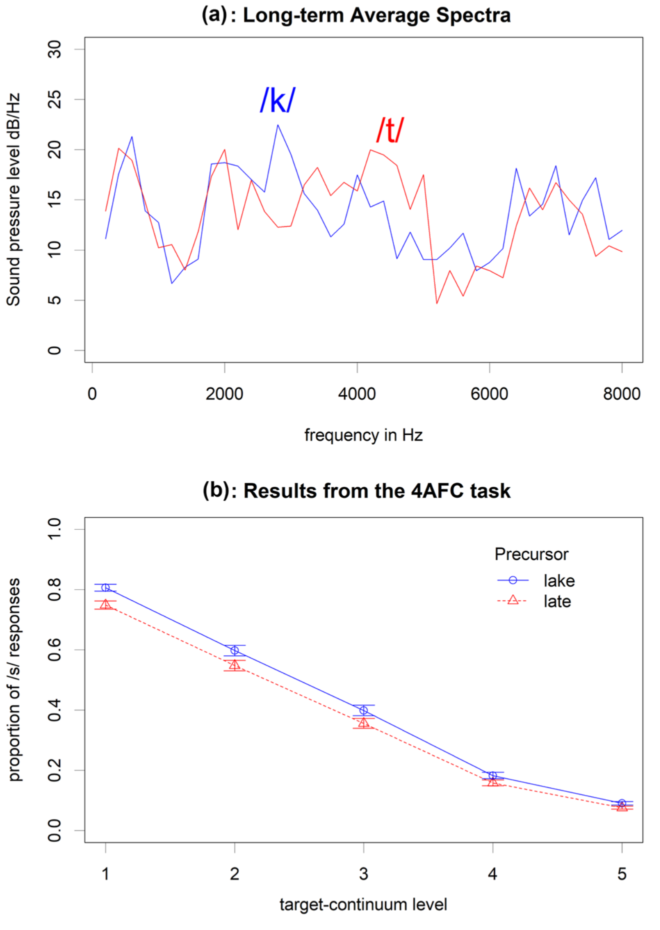

Fig 2. Panel A shows the spectra (bin width: 200 Hz) of the stop releases used in the /t/- and /k/-biasing contexts (*isolate* and *maniac*) in Luthra et al. (2021). Panel B shows the results of the cross-splicing experiment indicating CfC effects caused by these stop releases in a lexically unbiased context ([leɪ$^{k}/_{t}$], which can be interpreted as *lake* or *late*). Error bars are based on the standard error of the regression coefficient for context in the generalized linear mixed-effects model.

trimmed nonword-nonword contexts confirm that the acoustic differences had consequences for perceptual identification of those sounds (44% vs. 34% /t/ responses for the stops selected for use in the *isolate* and *maniac* contexts, respectively; 66% vs. 38% /l/ responses for the liquids selected for use in the *questionnaire* and *pocketful* contexts, respectively). One might argue that these data could be used to predict differences in CfC arising from these sounds. Identification rates of context sounds, however, are not reliable predictors of CfC. In a classic demonstration of this, Mann (1986) showed that the strength of CfC effects in English liquid-stop (i.e., [l/r]-[d/g]) sequences did not differ as a function of whether the (Japanese) listeners were at ceiling or at floor in identifying the liquids as /l/ or /r/. The dissociations in Pitt and McQueen (1998) and McQueen et al. (2009) discussed above are further demonstrations of this: A Ganong effect on the word-final fricatives does not necessarily predict a compensation effect on the following word-initial stops. What is required, therefore, is a direct test of the CfC effect arising from the different ambiguous sounds.

To this end, a cross-splicing experiment tested the hypothesis that the acoustic difference between the final ambiguous stops in the *isolate-maniac* pair of context words would result in CfC in the same direction as predicted by LCfC. Materials, data, and analysis files are available at: https://osf.io/pke94. The ambiguous stops, more [t]-like (alveolar) in the *isolate* context, more [k]-like (velar) in the *maniac* context, were spliced into a lexically unbiased [leɪ$^{k}/_{t}$] context (*late* and *lake* are both words) taken from the *isolate-isolake* continuum. They were followed by the Luthra et al. (2021) target stimuli (steps on [s]-[ʃ] word-word continua, e.g., *same-shame*). Participants performed a 4AFC task, categorizing both words in the stimulus sequences (e.g., as *late* or *lake* and as *same* or *shame*). If the acoustic difference in the context stops is strong enough to induce CfC, there should be fewer alveolar (/s/) responses after [leɪ$^{k}/_{t}$] derived from *isolate* (with a more alveolar release) than after [leɪ$^{k}/_{t}$] derived from *maniac* (with a less alveolar release).

Ideally, we would have carried out an equivalent cross-splicing experiment based on the *pocketful/questionnaire* stimuli. As noted above, however, the voicing difference (see Fig. 1) results in qualitatively different liquid sounds. This makes it difficult to compare the sounds and to make strong predictions about the spectral contrast effects that might arise from their acoustic properties. It also rules out a cross-splicing experiment because the strong vowel-liquid coarticulation makes it impossible to splice the two liquids onto the same vowel without creating acoustic discontinuities.

3.1. Method

3.1.1. Participants

We recruited 31 self-identified native speakers of English residing in the United States from the platform prolific.co within an age range from 18 to 40. The final sample had a median age of 28, ranging from 19 to 40. The task was advertised as taking 25 min and the median completion time was just below 22 min. Participants were rewarded with about 8.50 GBP per hour. The research was performed in accordance with the rules and regulations of the University of Malta Research Ethics committee.

3.1.2. Stimuli

For the precursor stimuli, we used the stop releases from the ambiguous stimuli from the *isolate-isolake* continuum and from the *maniat-maniac* continuum and spliced them onto the syllable [leɪ] from the *isolate-isolake* continuum, since this led to an easy answer format for participants (i.e., is the first part *lake* or *late*?). For simplicity, we will call the ambiguous stimulus from the [t]-biasing context (*isolate*) “late”, and the ambiguous stimulus from the [k]-biasing context (*maniac*) “lake”.

For the target stimuli, we used the original targets from five different [s]-[ʃ] continua: *same-shame*, *sell-shell*, *sign-shine*, *sip-ship*, and *sort-short*. We adjusted the peak loudness of these target words to match the maximal loudness in the *lake-late* stimuli. Note that this was necessary because otherwise the precursor, which was based on an unstressed syllable, would be perceived as having less prosodic weight which in turn would make compensation for coarticulation less likely (Kim, Mitterer, & Cho, 2018; Kuzla, Ernestus, & Mitterer, 2010). Given the five target continua had five steps each, this gives rise to a total of 50 stimuli (5 continua times 5 steps times 2 precursors).

3.1.3. Procedure

The experiment was run online using jsPsych (de Leeuw, 2015) as a four-alternative forced-choice task. Before the main experiment, participants were asked to rate the retail price of their headphones and then tested whether they used headphones by presenting stimuli with inter-aural time differences that had to be located in virtual space. After this task, which took about 1 min, they were instructed that they would hear a two-word sequence and that they had to identify both words by clicking on one of four tiles presented in the center of the screen. The upper two tiles always contained *late* and the lower two *lake*. The [s]-initial words were always presented on the right two tiles, while the [ʃ]-initial words appeared on the left two tiles. Participants were instructed to click on the tile that contained the two-word sequence they heard. Each of the 50 stimuli was presented six times, leading to 300 trials in total. Every 40 trials participants were given feedback on how many trials they had already completed and invited to take a short break.

3.2. Results

The data from one participant were removed because the participant failed the headphone test (only 12 out of 20 trials were correct). The data from the remaining participants show that they did not identify the precursors differently (*late*-stimulus: 52.1% *late* responses, *lake*-stimulus: 50.7% *late* responses). The data were analyzed with generalized linear mixed-effects models using a binomial/logit link function, with participant as a random effect and with all possible random slopes specified (though correlations between random slopes were removed to achieve convergence). For the perception of the precursor syllable, the source ([k]-biasing or [t]-biasing) did not lead to a difference in the identification as *late* or *lake* (regression weight for precursor: $b = 0.348$, $SE = 0.2807$, $z = 1.241$, $p = .215$). There was a large range of response patterns for the precursors, with some participants nearly exclusively perceiving them as *lake*, others exclusively as *late*, and others producing mixed responses with differences in both directions (i.e., either more *late* responses to the *lake* stimulus or to the *late* stimulus).

For the target stimuli, the model, with both continuum step and precursor as contrast-coded predictors, showed a significant effect for both predictors (see Fig. 2 and Table 1). The effect of the precursor is in the direction of CfC, with fewer alveolar responses if the preceding stimulus was "alveolar." Note that "alveolar" here means that the stop release was taken from the [t]-biasing context, even though there was no lexical bias here, only the acoustic difference. That is, even without a lexically biasing context, these stimuli are acoustically sufficiently different to trigger CfC effects in the direction predicted by the lexically biasing contexts used by Luthra et al. (2021).

Table 1
Results from the generalized linear mixed-effects model for fricative responses on the target continuum

	B (*SE*)	*z*	*p*
Intercept	−0.758 (0.270)	−2.812	.005
Precursor	−0.301 (0.072)	−4.187	<.001
Target continuum	−1.231 (0.061)	−20.129	<.001

4. Discussion

Luthra et al. (2021) present apparent evidence of LCfC for two sets of stimuli. For the set based on the context word pair *isolate-maniac*, we have presented evidence that an acoustic confound (different word-final stops) may underlie the apparent lexical effect. Note that the differences in CfC with these stimuli arose despite the lack of clear differences in their identification as /t/ or /k/. This underlines the point that we made earlier: that identification data are not sufficient to estimate the CfC potential of a set of stimuli. One might argue, however, that the apparent LCfC effect in the target article is larger than that reported here and hence that at least a component of the original effect must be due to the lexical bias manipulation. But it is possible that different effects were amplified when they were combined in the original study. That is, the use of ambiguous stimuli with acoustic differences combined with the use of words with TP biases could have increased the chances of learning based on experiment-internal bias. As different acoustic signals were used in different biasing contexts with different TP biases, it was easier for participants to learn the statistical dependencies in the experiment (e.g., one ambiguous sound always comes after *isola-*, and another ambiguous sound always comes after *mania-*). The original result could thus be the result of acoustic effects, TP or experiment-induced biases, or their combination, and, when combined, those effects could have amplified each other. A true LCfC effect with these stimuli thus remains to be shown in an experiment without an acoustic confound and with TPs and experimental biases controlled.

We were not able to test for CfC arising from the acoustic differences in the stimulus set based on the context word pair *pocketful-questionnaire*. We, therefore, do not know whether there is or is not a confound in these materials. Note, however, that the difference in acoustic energy in the range of the second and third formants could make the lower parts of the fricative noise more prominent after the devoiced *pocketful* than after the voiced *questionnaire*. This would lead to more /s/ responses after *questionnaire*, in line with the apparent LCfC effect. A similar argument to that for *isolate-maniac* can, therefore, be made: A true LCfC effect with the *pocketful-questionnaire* stimuli needs to be demonstrated after this possible acoustic confound has been removed, and again in a situation with TPs and experimental biases controlled. As previously noted, the pilot liquid identification data should not be used to predict what the acoustically based CfC might look like for *pocketful-questionnaire*.

4.1. *The fragility of the paradigm*

Luthra et al. (2021) confirm just how fragile the underlying effects are in the LCfC paradigm (e.g., only 20 of the 130 pairs that they piloted gave a CfC effect). This fragility may have arisen in part because of the intonation of the initial recordings used to make the stimuli (see Supplementary Materials). McQueen et al. (2009) suggested it might be time to stop using the paradigm given not only this fragility but also the need to control for so many factors. Unfortunately, however, the paradigm is exactly what is needed for precise claims to be made about the cognitive mechanisms underlying context effects in speech comprehension. It is thus probably necessary to continue to use the paradigm (or variants thereof). But it needs to be used very carefully. This point holds true across cognitive science: the devil and the angels are in the details!

4.2. *The LCfC paradigm*

If the paradigm is still to be used, the next question to ask is: what is the best way to use it? Luthra et al. (2021) argue for the inclusion of pretests to establish that the stimuli are capable of generating Ganong and CfC effects. We agree that these procedures are important. If the lexical bias is too weak to generate a Ganong effect and/or the acoustics of the stimuli are not conducive for CfC, as indeed was the case in the majority of the stimulus sets that Luthra et al. started out with, then there is no reason to expect LCfC. But pretesting is not the only way to ensure that the conditions are appropriate to observe LCfC. An alternative (adopted by Magnuson et al., 2003, McQueen et al., 2009, and Pitt & McQueen, 1998) is to use a 4AFC task in the main experiment (this allows the researcher to check that there is a Ganong effect on categorization of the ambiguous sounds at the end of the first words in the test sequences) and additional trials/conditions (other than the critical LCfC trials) that provide an independent measure that the stimuli can generate CfC. We suggest that this alternative is preferable to that proposed by Luthra et al. because it creates the opportunity to test for possible dissociations of the two basic effects, as discussed in Section 2.1. The best approach, however, would be a combination: first pretesting (else the researcher takes a very large risk) and second the use of the 4AFC task in the main experiment.

4.3. *The feedback debate*

It has long been known that listeners use context and prior knowledge in speech comprehension (Miller, Heise, & Lichten, 1951). The question addressed in the LCfC literature is how they do so. This remains a key question in speech science, and, generalizing to other domains of perception, in cognitive science more broadly. To answer it, distinctions need to be made between different cognitive mechanisms, for example, the distinction between online feedback and feedback for perceptual learning (Norris et al., 2003). As already discussed, it has been shown that learning processes can be responsible for apparent LCfC effects. Feedback could also play a role in attention or binding (Norris, McQueen, & Cutler, 2016). Answering the question whether or not there is online feedback in different perceptual domains (i.e., not only in speech recognition) thus requires careful analysis of whether apparent evidence of online feedback might be due to other mechanisms.

Another aspect of this question concerns the function that online feedback might serve. It has been argued, for speech perception, that online feedback cannot improve the performance of an optimal recognizer and thus is not necessary (Norris, McQueen, & Cutler, 2000). Although Magnuson, Mirman, Luthra, Strauss, and Harris (2018) have argued that online feedback can improve the recognition performance of the TRACE model (McClelland & Elman, 1986), this is because the interactive-activation framework of TRACE is not optimal (Norris, McQueen, & Cutler, 2018). It has thus not been shown that online feedback is beneficial across all types of word-recognition models, and hence its implementation-independent function has not yet been specified. Again, across domains of cognitive science, it will be necessary not only to search for evidence that can unambiguously be attributed to online, top-down feedback, but also to specify the function of this process.

5. Conclusion

Luthra et al.'s (2021) claim of a robust LCfC effect is premature. The effect they observe could be due to TP or experiment-induced biases, acoustic effects, or a combination of these factors, rather than to online, top-down feedback from the lexicon to earlier stages of perceptual processing. Given this, and that Luthra et al. motivate their study on the grounds that the prior literature on LCfC was inconsistent, the LCfC paradigm has thus not yet made a convincing case for online lexical feedback. In contrast, the paradigm has instead made a strong case against this kind of feedback (McQueen et al., 2009). That case was further strengthened by the BF analyses we report here. It nevertheless remains true that the paradigm has the potential to generate evidence of online top-down processing. But that evidence remains as elusive as a white Christmas.

Acknowledgments

We thank Jeff Starns for his advice on Bayesian statistics. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2020S1A5A2A03042969).

References

Bushong, W., & Jaeger, T. F. (2019). Dynamic re-weighting of acoustic and contextual cues in spoken word recognition. *Journal of the Acoustical Society of America*, *146*, EL135. https://doi.org/10.1121/1.5119271

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(781), 1–17. https://doi.org/10.3389/fpsyg.2014.00781

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. https://doi.org/10.1016/0749-596X(88)90071-X

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Kim, S., Mitterer, H., & Cho, T. (2018). A time course of prosodic modulation in phonological inferencing: The case of Korean post-obstruent tensing. *PLOS ONE*, *13*(8), e0202912. https://doi.org/10.1371/journal.pone.0202912

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*, 1–15.

Kuzla, C., Ernestus, M., & Mitterer, H. (2010). Compensation for assimilatory devoicing and prosodic structure in German fricative perception. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory phonology* (Vol. *10*, pp. 731–758). Mouton.

Luthra, S., Peraza-Santiago, G., Beeson, K., Saltzman, D., Crinnion, A. M., & Magnuson, J. S. (2021). Robust lexically-mediated compensation for coarticulation: Christmash time is here again. *Cognitive Science*, *45*, e12962. https://doi.org/10.1111/cogs.12962

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.

Magnuson, J. S., Mirman, D., Luthra, S., Strauss, T., & Harris, H. D. (2018). Interaction in spoken word recognition models: Feedback helps. *Frontiers in Psychology*, *9*, 369. https://doi.org/10.3389/fpsyg.2018.00369

Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r". *Cognition*, *24*, 169–196.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

McQueen, J. M. (2003). The ghost of Christmas future: Didn't scrooge learn to be good? Commentary on Magnuson, McMurray, Tanenhaus and Aslin (2003). *Cognitive Science*, *27*, 795–799. https://doi.org/10.1207/s15516709cog2705_6

McQueen, J. M., Jesse, A., & Norris, D. (2009). No lexical–prelexical feedback during speech perception or: Is it time to stop playing those Christmas tapes? *Journal of Memory and Language*, *61*, 1–18. https://doi.org/10.1016/j.jml.2009.03.002

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, *41*, 329–335. https://doi.org/10.1037/H0062491

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*, 299–325. https://doi.org/10.1017/S0140525×00003241

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*, 204–238. https://doi.org/10.1016/S0010-0285(03)00006-9

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience*, *31*, 4–18. https://doi.org/10.1080/23273798.2015.1081703

Norris, D., McQueen, J. M., & Cutler, A. (2018). Commentary on "Interaction in spoken word recognition models." *Frontiers in Psychology*, *9*, 1568. https://doi.org/10.3389/fpsyg.2018.01568

Pitt, M. A., & McQueen, J. M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, *39*, 347–370. https://doi.org/10.1006/jmla.1998.2571

Poellmann, K., McQueen, J. M., & Mitterer, H. (2011). The time course of perceptual learning. In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011* [ICPhS XVII] (pp. 1618–1621). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.

Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.), *Categorical perception* (pp. 89–112). Cambridge, England: Cambridge University Press.

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.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1: Bayes Factor calculations for Experiment 3 in McQueen et al. (2009).