Robust lexically-mediated compensation for coarticulation: Christmash time is here again

Sahil Luthra[1], Giovanni Peraza-Santiago[1], Keia'na Beeson[1], David Saltzman[1], Anne Marie
Crinnion[1], & James S. Magnuson[1,2,3]

[1] University of Connecticut

[2] BCBL. Basque Center on Cognition Brain and Language, Donostia-San Sebastián, Spain

[3] Ikerbasque. Basque Foundation for Science, Bilbao, Spain

Author Note

CORRESPONDING AUTHOR:
        Sahil Luthra
        Psychological Sciences
        University of Connecticut
        Storrs, CT 06269-1020
        sahil.luthra@uconn.edu

Abstract

A long-standing question in cognitive science is how high-level knowledge is integrated with sensory input. For example, listeners can leverage lexical knowledge to interpret an ambiguous speech sound, but do such effects reflect direct top-down influences on perception or merely post-perceptual biases? A critical test case in the domain of spoken word recognition is *lexically-mediated compensation for coarticulation* (LCfC). Previous LCfC studies have shown that a lexically-restored context phoneme (e.g., /s/ in *Christma#*) can alter the perceived place of articulation of a subsequent target phoneme (e.g., the initial phoneme of a stimulus from a *tapes-capes* continuum), consistent with the influence of an unambiguous context phoneme in the same position. Because this phoneme-to-phoneme compensation for coarticulation is considered sublexical, scientists agree that evidence for LCfC would constitute strong support for top-down interaction. However, results from previous LCfC studies have been inconsistent, and positive effects have often been small. Here, we conducted extensive piloting of stimuli prior to testing for LCfC. Specifically, we ensured that context items elicited robust phoneme restoration (e.g., that the final phoneme of *Christma#* was reliably identified as /s/) and that unambiguous context-final segments (e.g., a clear /s/ at the end of *Christmas*) drove reliable compensation for coarticulation for a subsequent target phoneme. We observed robust LCfC in a well-powered, pre-registered experiment with these pretested items (N=40) as well as in a direct replication study (N=40). These results provide strong evidence in favor of computational models of spoken word recognition that include top-down feedback.

1. Introduction

Cognitive scientists have long observed that an individual's interpretation of sensory information can be shaped by context. For instance, an individual will report hearing a speech sound that has been replaced by noise (e.g., hearing the phoneme /s/ in the frame *legi_lature*, where the critical phoneme has been replaced by a cough; Warren, 1970), and an individual's estimate of an object's size can be influenced by the width between a person's hands (Stefanucci & Geuss, 2009). While such context effects are ubiquitous across a range of domains in cognitive psychology, an ongoing debate – whether in the domain of language (e.g., Magnuson, Mirman, Luthra, Strauss, & Harris, 2018; Norris, McQueen, & Cutler, 2018) or the domain of vision (e.g., Firestone & Scholl, 2014, 2015; Gilbert & Li, 2013; Lupyan, Abdel Rahman, Boroditsky, & Clark, in press; Schnall, 2017a, 2017b) – centers on how contextual information is integrated with sensory signals. In particular, do contextual effects on sensory processing reflect influences on perception itself, or does context only affect an individual's post-perceptual decisions?

The nature of this debate can be better understood by considering computational models of cognition. In these models, processing is often conceptualized as occurring in a hierarchy, with lower levels corresponding to early stages of processing; in speech perception, this might be the processing of acoustic-phonetic information, and in vision, this might be the processing of low-level features like brightness and line orientation. Higher levels correspond to the processing of more complex information (e.g., word-level information in spoken word recognition, or object recognition in visual processing). The ongoing debate over context effects on perception relates to information flow within such hierarchies and how competing models explain context effects. In *interactive* models, information is fed back from higher levels of processing to directly affect the activity of lower levels. Contrastively, *autonomous* models do not allow for feedback from higher

to lower levels; instead, information from different levels of perceptual processing is combined post-perceptually. This debate has been particularly difficult to resolve in the domain of spoken language processing, which is our focus.

To illustrate, consider the finding that a listener's interpretation of an ambiguous speech sound can be influenced by lexical knowledge (e.g., Ganong, 1980). For instance, listeners are more likely to interpret a speech sound that is ambiguous between /g/ and /k/ as *g* if the sound is followed by *-ift*, but as *k* if the speech sound is followed by *-iss*. Interactive models of spoken word recognition, such as the TRACE model (McClelland & Elman, 1986), account for this effect by allowing activation to flow from a higher (lexical) level of processing to influence processing at a lower (phonological) level. From this perspective, lexical knowledge can feed back to restore an ambiguous speech sound in a literal top-down fashion. By contrast, strictly feedforward models, such as Merge (Norris, McQueen, & Cutler, 2000) and Shortlist B (Norris & McQueen, 2008), explain this effect by allowing lexical and phonological information to be combined only post-perceptually (e.g., via a second set of phoneme "decision nodes" in Merge that receive input from both "perceptual" phoneme nodes and lexical nodes). Proponents of interactive models have noted a correspondence between the idea of feedback in computational models and the documented presence of feedback loops in neurobiology (Magnuson et al., 2018; Montant, 2000); neural feedback has been argued to help individuals build a stable (if not entirely veridical) representation of the world, despite the variability present in sensory input (Gilbert & Li, 2013). Critics of interactive models argue that the presence of feedback connections in neurobiological models should not count as evidence for feedback in computational models, as the function of recurrent neural connections in the brain has not been adequately characterized (Norris, McQueen, & Cutler, 2016). Furthermore, they have expressed concern that if higher levels of processing directly were

to modulate lower levels of processing, individuals' perception would reflect hallucinations rather than the actual state of the world (Norris et al., 2016).

How might we distinguish between interactive and autonomous accounts if both provide explanations for lexical context effects? Elman and McClelland (1988) proposed a particularly clever test. First, they noted that the phenomenon of *compensation for coarticulation* appears to have a clearly prelexical, perceptual locus. When a speaker must produce a sound with a front place of articulation (e.g., /s/) and then one with a back place of articulation (e.g., /k/), or vice-versa, coarticulation (motoric constraints on articulation) may keep them from reaching the canonical *place of articulation* (henceforth, PoA) for the second sound, leading to an ambiguous production. Thus, when listeners hear ambiguous tokens after a sound with a front PoA, they appear to compensate for coarticulation and attribute ambiguity to the speaker trying to reach a relatively distant, posterior PoA (Mann & Repp, 1981; Repp & Mann, 1981, 1982). Second, Elman and McClelland proposed that if lexical information directly modulates perception by sending top-down feedback to phoneme-level representations, then a lexically-restored phoneme could induce compensation for coarticulation, providing strong evidence for an interactive model.

To test whether a lexically-restored phoneme could drive compensation for coarticulation, Elman and McClelland (1988) presented listeners with two consecutive stimuli. The first was a *context* item that ended with a speech sound that was ambiguous between /s/ and /ʃ/ ("sh," with back PoA); critically, this ambiguous speech sound was embedded in a context where lexical knowledge could disambiguate the intended phoneme (e.g., *Christma#* or *fooli#*, where # denotes the same ambiguous sound). The second stimulus was a *target* item from a *tapes-capes* continuum (where /t/ has a front PoA and /k/ has a back PoA). Critically, the authors observed that lexical context influenced compensation for coarticulation. When lexical information guided the listener

to interpret the ambiguous sound # as /s/, participants were more likely to hear the subsequent sound as /k/, just as would be the case if the context word had ended with a clear /s/. Similarly, when lexical context restored the # to /ʃ/, compensation for coarticulation was observed in the same way as if the context item had ended with a clear /ʃ/. This finding was particularly momentous because it seemed that the *lexically-mediated compensation for coarticulation (LCfC)* effect could only be produced by an interactive mechanism. As such, this logic has been described as the "gold standard" for evaluating claims that there is feedback in spoken word processing, even by proponents of strictly feedforward models (p. 9; Norris et al., 2016), since it tests whether lexical knowledge can induce a process that is agreed to occur at a prelexical stage.

Nevertheless, results from other LCfC studies have been inconsistent. Despite some replications of the original finding (Magnuson, McMurray, Tanenhaus, & Aslin, 2003a; Samuel & Pitt, 2003), there are some reasons to be skeptical. For instance, Pitt and McQueen (1998) postulated that the original results might have been driven by diphone transitional probabilities rather than lexical information, possibly consistent with either an autonomous or interactive architecture. In a study using nonword contexts, they observed CfC patterns consistent with transitional probabilities (in the absence of lexical context) and failed to observe LCfC when transitional probabilities were equi-biased. However, Magnuson et al. (2003a) observed an LCfC effect even when diphone transitional probabilities had opposite biases as lexical contexts. Subsequently, McQueen, Jesse, and Norris (2009) suggested that the LCfC effects observed by Magnuson et al. (2003a) might specifically be attributable to the fact that listeners in that experiment heard contexts with clearly produced, lexically consistent endings (e.g., *Christmas*) and contexts with ambiguous word-final phonemes (*Christma#*), but never contexts with lexically inconsistent endings (*Christmash*). Thus, a listener's inclination to interpret the final phoneme in

a lexically consistent way may have been because of an experiment-induced bias, rather than because of lexical knowledge *per se*. Both of these concerns must be addressed in LCfC studies attempting to distinguish between autonomous and interactive accounts.

Although the scorecard favors interaction (approximately 60% of reported results have been positive), there could also be a file drawer problem. Significant results might be false-positive flukes, while additional replication failures may have been set aside in a metaphorical filing cabinet. In considering this possibility and reviewing the prior literature, we also noted a salient gap in many studies. Few have included pretests of context and/or target materials to establish that context items actually provide the conditions necessary for observing Ganong (1980) effects, or that target items provide conditions necessary for observing CfC with unambiguous preceding context phonemes. In the current study, we set out to test LCfC only after establishing that we could detect Ganong effects with candidate context items and CfC effects with candidate target items. After observing apparently robust LCfC in an initial pre-registered experiment, we conducted our own direct replication – and again observed robust LCfC. Critically, in both studies, listeners only heard context items that ended in ambiguous phonemes, allowing us to address the concern of McQueen et al. (2009). After presenting the experiment and replication, we demonstrate that we can account for the pattern of results across LCfC studies by appealing to lexical status but not to transitional probabilities, allowing us to address the concern of Pitt and McQueen (1998). We close with a discussion of the methodological and theoretical implications of our results.

2. Piloting

To ensure that our materials would provide conditions capable of detecting LCfC, we performed extensive piloting to select context items on which strong phoneme restoration effects could be observed, and target items on which strong CfC effects could be observed when preceded by unambiguous context segments. If context and target items do not meet these preconditions, there would be no reason to expect to observe *L*CfC.

Potential context items ended with a speech sound that had an ambiguous place of articulation. These items were created by morphing a word and a nonword, one with a front place of articulation and one with a back place of articulation. In constructing context items, we required that the lexical endpoints be high-frequency words, and we attempted to maximize the number of syllables so that the final phoneme would have strong lexical support. The word-final phoneme could not appear anywhere else in the word.

Potential target continua were constructed by morphing two minimally contrastive words, one of which had a front place of articulation and one of which had a back place of articulation. As before, we required that all continuum words had a high lexical frequency.

All items were recorded by a male native speaker of American English who produced each item multiple times. We used STRAIGHT (Kawahara et al., 2008) to create 11-step continua for context and target stimuli; STRAIGHT requires the experimenter to identify landmarks in both the temporal and spectral domains prior to interpolation, and the resultant continua sound more naturalistic than continua produced by waveform averaging.

We conducted three pilot studies to validate our stimuli. All studies were conducted using the online experiment builder Gorilla (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020) and the online data collection platform Prolific (https://www.prolific.sc). All participants were adult native speakers of English with no reported history of speech, language, hearing or

vision impairments. Participants were paid at the Connecticut minimum wage ($11/hour at the time) based on the estimated time to complete the study, and each participant was only eligible to participate in one pilot study or experiment. All procedures were approved by the University of Connecticut Institutional Review Board (IRB), and participants provided informed consent before beginning an experiment. Participants answered a set of demographics questions and completed a "headphone screening" prior to the study (a stereo listening test that is virtually impossible to pass without headphones; Woods, Siegel, Traer, & McDermott, 2017). If participants failed the headphone check twice, their data were excluded from analyses.

The first pilot study (N=40) was used to select context continua. We began with a set of 35 potential context items, which ended with either /d/ or /g/ (e.g., *episode/*episogue*), /l/ or /r/ (*questionnail/questionnaire*), /s/ or /ʃ/ (*aboliss/abolish*), or /t/ or /k/ (*isolate/*isolake*). Because our paradigm requires listeners to leverage the beginning of a context word (e.g., *aboli-*) to interpret a word-final ambiguous speech sound (e.g., s/sh), we wanted to ensure that for our context words, subjects' interpretation of the word-final segment was guided by the preceding context. One group of subjects (n=20) heard the full stimuli, such that they made judgments on word-nonword continua (e.g., *abolish-*aboliss*). Another group (n=20) heard trimmed items where lexical information was removed, such that they heard nonword-nonword continua (e.g., *abolish-*aboliss* were trimmed to *ish-iss*).[1] For both groups, listeners heard 9 steps (steps 2-10) from every continuum and were asked to indicate what the final phoneme of the stimulus was. Each participant heard each stimulus twice. Stimuli were blocked by contrast (e.g., a given subject may have heard

---

[1] To create these nonword-nonword continua, we identified *cutpoints* for each unambiguous stimulus (i.e., timepoints where cuts could be made to each endpoint step to yield the desired nonword); stimuli were cut at the zero-crossing closest to the cutpoint. For intermediate steps along each continuum, the cutpoint was identified using linear interpolation between the cutpoints of the endpoint steps.

all the items from the /d/-/g/ continua before hearing any of the /t/-/k/ stimuli), with random ordering of items within blocks, and the order of contrasts was counterbalanced using a Latin square. For a given subject, the left response button either corresponded to all the front responses (/s/, /t/, /d/, /l/) or all the back responses (/ʃ/, /k/, /g/, /r/), with the specific mapping counterbalanced across participants. In order for a context continuum to be selected for future piloting, there had to be at least one ambiguous step with a lexical effect – that is, there had to be a step for which subjects who heard the full stimuli (i.e., a word-nonword continuum, such as *abolish-*aboliss*) made more lexically consistent responses than subjects who heard the trimmed stimuli (i.e., a nonword-nonword continuum). In this way, we selected 10 continua for further piloting.

In a second pilot (N=20), we sought to validate our target stimuli. We began with a set of 32 target continua, each of which *began* with /d/ or /g/ (e.g., *deer*/*gear*), /l/ or /r/ (*lake*/*rake*), /s/ or /ʃ/ (*same*/*shame*), or /t/ or /k/ (*tea*/*key*). For this pilot, subjects heard 9 steps (steps 2-10) from each continuum and were asked to identify the first phoneme of each item. As before, each participant heard each stimulus twice. Stimuli were blocked by contrast (with contrast order counterbalanced using a Latin square and order randomized within blocks), and response mappings were counterbalanced across participants. Piloting yielded five usable continua per contrast (i.e., continua where one endpoint elicited a front response, one endpoint elicited a back response, and the identification function resembled a sigmoidal curve). Through this pilot, we identified the most ambiguous step for each of the 20 selected continua. We selected this maximally ambiguous step as well as the two steps on either side of it for use in a final pilot study.

In the final pilot study (N=20), we sought to establish that our stimuli could elicit robust compensation for coarticulation when listeners heard *unambiguous* context items prior to the target continuum (in contrast to the main experiment, in which context items would end with ambiguous

speech sounds). In this pilot study, listeners heard the context words selected from the first pilot (e.g., a clear, lexically consistent production of *abolish* or *dangerous*) prior to the five continuum steps selected from the second pilot, with the constraint that the same phoneme contrast could not be used for both the context-final phoneme and the target-initial phoneme (e.g., context items ending with /d/-/g/ could not be paired with target items beginning with /d/-/g/). Context items and target items were all scaled to 70dB SPL prior to concatenation. Each of the 130 context-target pairs was heard once by each participant. As before, stimuli were blocked by contrast (with contrast order counterbalanced using a Latin square, and order randomized within blocks), and response mappings were counterbalanced across participants. We observed robust compensation for coarticulation (e.g., more *front* responses on ambiguous steps of target continua when the preceding context word had ended with a segment that had a *back* place-of-articulation) for 20 context-target pairs, and these stimuli were used in the main experiment.

It is striking that we observed compensation for coarticulation for 20 of the 130 context-target pairs we tested in this pilot. Moreover, while all our target stimuli beginning with fricatives (/s/, /ʃ/) survived this final pilot, we did not observe robust compensation for coarticulation for target items beginning with other phonemes (/d/, /g/, /l/, /r/, /t/, /k/). It is also notable that only 10 of 35 candidate context items survived our pretesting. We return to this issue in the General Discussion.

3. Experiment

Following piloting, we conducted a well-powered, pre-registered study to test for LCfC. Critically, participants in the main LCfC study only heard ambiguous phonemes at the end of context items, allowing us to address the potential confound described by McQueen et al. (2009).

3.1. Methods

3.1.1. Materials. To select stimuli for this experiment, we conducted several pilot studies, and only stimuli that showed the expected effects in all pilots were included. We selected a set of four validated context items (from an initial set of 35): *isolate/*isolake*, *maniac/*maniat*, *pocketful/*pocketfur*, and *questionnaire/*questionnail*. Based on the results of our piloting, we selected five target continua (from an initial set of 32) for use in the main experiment: *same-shame*, *sell-shell*, *sign-shine*, *sip-ship*, and *sort-short*.

3.1.2. Procedure. After conducting the pilot studies described in Section 2, we pre-registered the procedure for our main experiment on the Open Science Framework. As with the pilot studies, the experiment was implemented using Gorilla and data were collected through the Prolific platform. Participants provided informed consent, answered demographics questions, and completed the headphone screening test prior to beginning the experiment.

Each trial consisted of an *ambiguous* context item (e.g., *pocketfu#*) immediately followed by a target word. (Note that is in contrast to the third pilot described in Section 2, where the context stimulus ended with an unambiguous phoneme.) Participants had to decide whether the target word they heard began with 's' or 'sh' by pressing the appropriate key. Subjects completed 6 blocks of this task, with each block consisting of all 100 possible trials (4 context items x 5 target continua x 5 steps / target continuum) in random order. Response mappings (i.e., whether the 's' button response was on the left or on the right) were counterbalanced across participants. In total, the experiment took approximately 45 minutes to complete.

3.1.3. Participants. Sixty-one participants were recruited for this experiment. All individuals recruited were aged 18-34 (due to expected age-related declines in auditory acuity after

11

age 35) and self-reported being native speakers of English with normal or corrected-to-normal vision, normal hearing, and no history of speech or language impairments. Participants were paid $8.25 for their participation, consistent with Connecticut minimum wage ($11/hour at the time of the study).

We decided *a priori* to exclude participants who failed the headphone checks, had low (<80%) accuracy in their classification of the clear endpoints of the target continua, and/or failed to respond to 10% or more of trials. (Trials timed out after 6 seconds.) This led to a final sample size of 40 (22 female, 18 male), with participants ranging in age from 18 to 34 (mean: 27). Note that because effect sizes in LCfC experiments are known to be small, we decided to use a larger sample size than in previous studies (Elman & McClelland, 1988; Magnuson et al., 2003a; Pitt & McQueen, 1998), which have used sample sizes between 16 and 30.

3.2. Results and Discussion

We observed robust lexically-mediated compensation for coarticulation in our experiment. As shown in Figure 1, participants were more likely to indicate that a target item began with a front PoA if the context item had an implied back PoA (*maniac*, *questionnaire*) than if the context item had an implied front PoA (*isolate*, *pocketful*).
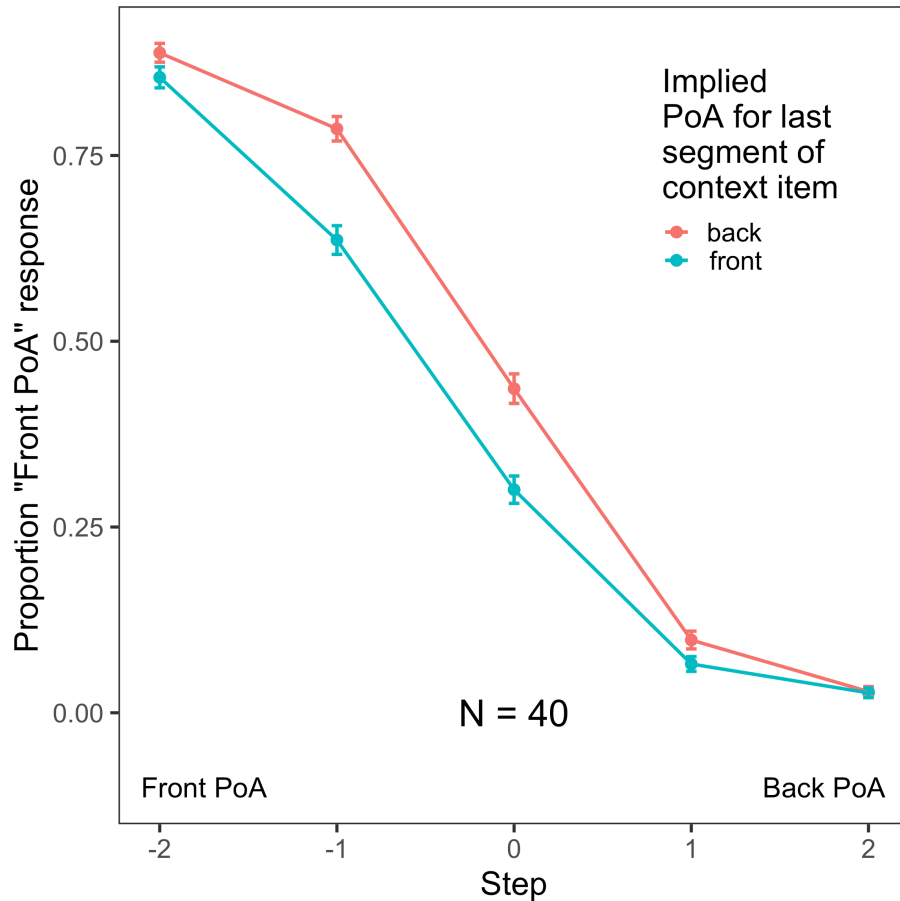
Figure 1. In this experiment, participants were asked to identify the first phoneme of a target item that was preceded by a context item. Context items ended with an ambiguous place of articulation (e.g., a blend of *isolate* and *isolake*), but lexical information could be used to guide interpretation of the final segment. Target items began with a segment from a /s/-/ʃ/ continuum (e.g., a *sign-shine* continuum). The x-axis indicates whether the target began with a front (/s/) or back (/ʃ/) place of articulation (PoA); a value of 0 indicates the most ambiguous step from that continuum, as determined in pilot testing. The y-axis indicates the proportion of responses with a front PoA (i.e., a /s/ response). Critically, responses differed depending on whether the lexically implied final segment of the context item had a front PoA (*isolate*, *pocketful*) or a back PoA (*maniac*, *questionnaire*). Error bars indicate 95% confidence intervals.

Data were analyzed in R (R Core Team, 2019) using a mixed effects logistic regression. Models were implemented using the *mixed* function in the "afex" package (Singmann, Bolker, Westfall, & Aust, 2018); this package provides a wrapper to the *glmer* function in the "lme4" package (Bates, Maechler, Bolker, & Walker, 2015), allowing results to be obtained in an

ANOVA-like format. We tested for fixed effects of Context (front, back; coded with a [1,-1] contrast) and Step (centered, with 5 steps from -2 to +2). Our model also included random by-subject slopes for Context and Step, as well as their interactions, and random by-subject intercepts. This is both the maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013) and the most parsimonious structure (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017), as simplifying the random effect structure led to significantly worse model fit. The significance of main effects and interactions was estimated through likelihood ratio tests.

Analyses indicated a significant effect of the implied place of articulation of the Context item (χ^2 = 42.26, p < 0.0001), corresponding to robust lexically-mediated compensation for coarticulation. Specifically, participants made front-PoA responses 45% of the time when the last segment of the context item had an implied back-PoA and 38% of the time when it had an implied front PoA. Note that this difference (7%) is comparable in size to the LCfC effects observed in previous studies (9% difference in Elman & McClelland [1988]; 6% difference in Magnuson et al. [2003a]). We also observed a significant effect of Step (χ^2 = 124.35, p < 0.0001), indicating that listeners made more /s/ responses for target items that were closer to the /s/ end of the continuum. We also observed a significant interaction between the factors (χ^2 = 4.54, p = 0.03), indicating that the effect of lexical information (Context) was not constant across Steps.

Overall, our results suggest that the place of articulation of a lexically-restored phoneme (e.g., the /t/ at the end of *isola#*) could induce compensation for coarticulation on a subsequent front-back place of articulation continuum. Such an effect can be explained naturally by models of spoken word recognition that allow for activation feedback from the lexical level to the phonological level. In the General Discussion, we will consider alternative explanations based on sublexical transitional probabilities. While positive results have been observed in several previous

studies (Elman & McClelland, 1988; Magnuson et al., 2003a; Samuel & Pitt, 2003), some studies have failed to observe LCfC (McQueen et al., 2009 [which included an attempt to directly replicate Magnuson et al., 2003a]; Pitt & McQueen, 1998). As such, we opted to conduct a direct replication of our findings with an independent sample.

4. Replication

4.1. Methods

4.1.1. Stimuli. The same stimuli were used as in the previous experiment.

4.1.2. Procedure. The procedure for the replication experiment was identical to the procedure of the original experiment.

4.1.3. Participants. Sixty-three participants were recruited following the same methods as in our original experiment; however, no individual who had participated in the original experiment was eligible to participate in the replication experiment. Two subjects' data were excluded due to technical failures (e.g., poor internet connection). As before, participants' data were excluded if they failed the headphone screening, had low (<80%) accuracy in their classification of unambiguous endpoint stimuli from the target continuum and/or if they failed to respond to 10% or more of the trials. This resulted in a final sample size of 40 (22 female, 17 male, 1 who preferred not to say) for our analyses. Participants in the final sample ranged in age from 18 to 32 (mean: 23).

4.2. Results and Discussion

Results from the replication study are shown in Figure 2.
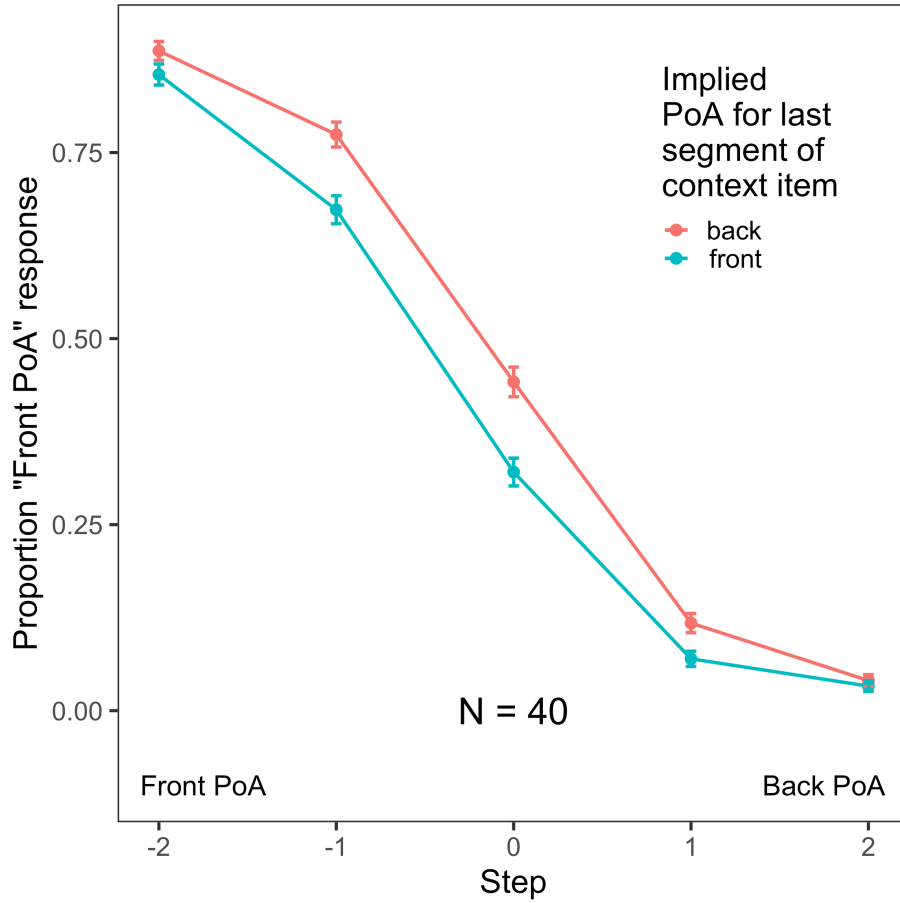
Figure 2. We replicated the findings from our initial experiment in a separate sample. The x-axis indicates whether the target began with a front (/s/) or back (/ʃ/) PoA; a value of 0 indicates the most ambiguous step from that continuum. The y-axis indicates the proportion of responses with a front PoA (i.e., a /s/ response). As before, we observed lexically-mediated compensation for coarticulation. Error bars indicate 95% confidence intervals.

Data were analyzed following the same procedure as above, and the same model structure was used as for the initial sample. As in the initial experiment, we obtained a significant effect of the implied place of articulation of the Context item ($\chi^2 = 39.28$, $p < 0.0001$), indicating robust lexically-mediated compensation for coarticulation in this independent sample. This corresponded to participants making front-PoA responses 45% of the time when the context item ended with an implied back-PoA segment and 39% of the time when the final segment of the context item had an implied front-PoA (a 6% difference, again comparable to previous studies). As before, we also

observed a significant effect of Step ($\chi^2 = 104.53$, $p < 0.0001$), indicating that listeners made more /s/ responses for target items that were closer to the /s/ end of the continuum. In contrast with the initial sample, however, we did not observe a significant interaction between these factors ($\chi^2 = 0.14$, $p = 0.71$). Nonetheless, we replicated the critical finding from our initial experiment in that we observed robust LCfC (indicated by the significant effect of Context).

4. General Discussion

A challenge for resolving the debate between autonomous and interactive models is that while lexically-mediated compensation for coarticulation (LCfC) is seen as the gold standard for evidence of feedback, LCfC effects have historically appeared to be rather fragile (McQueen et al. [2009], for instance, failed to replicate the Magnuson et al. [2003a] findings) and sensitive to a variety of parameters and conditions (Samuel & Pitt, 2003). In the current study, we therefore conducted extensive stimulus piloting before testing for LCfC. Specifically, we first established that we could observe lexical restoration of ambiguous phonemes at the ends of lexical context words (e.g., that an ambiguous fricative at the end of *Christma#* would be identified as /s/), and that we could observe compensation for coarticulation due to place of articulation in candidate target continua given unambiguous segments at the ends of context words (e.g., a clear /s/ at the end of *Christmas* or a clear /ʃ/ at the end of *foolish* would shift identification of steps in a following *tapes-capes* continuum). We observed robust LCfC when we combined items that survived our pilot criteria in our experiments: lexically-restored phonemes influenced perception of subsequent phonemes (e.g., a shift towards more *tapes* responses following *fooli#* vs. more *capes* responses following *Christma#*). Critically, participants in our experiment only ever heard ambiguous phonemes at the ends of context stimuli, suggesting that phoneme restoration was driven by the

lexicon rather than by experiment-induced biases (the explanation McQueen et al., [2009] proposed for the results of Magnuson et al. [2003a]). We observed LCfC effects both in an initial experiment *and* in a direct replication of the study, with effect sizes comparable to those reported by Elman and McClelland (1988) and Magnuson et al. (2003a). These results support interactive models of spoken word recognition, in which information is fed back from higher levels of processing to directly influence the activity of lower levels.

Pitt and McQueen (1998) have argued that results obtained in earlier LCfC studies may have been driven by diphone transitional probabilities rather than lexical information, potentially allowing positive results to be explained by models without feedback. That is, listeners may restore an ambiguous phoneme at the end of *Christma#* as /s/ not because of lexical knowledge but because the /s/ phoneme is more likely than /ʃ/ in this specific vowel context. Magnuson et al. (2003a) showed that LCfC effects are *not* attributable to transitional probabilities, observing LCfC even when lexical biases and transitional probabilities were in opposition. Furthermore, Magnuson, McMurray, Tanenhaus, & Aslin (2003b) reported not only that neither forward nor backward diphone transitional probabilities could explain all attested positive instances of LCfC in the literature, but also that there was no single *n*-phone that could explain all positive effects. We have updated that analysis to include the transitional probabilities of *every* context item used in any LCfC experiment of which we are aware (Appendix A). Initial results analyzing only the words used in our experiment (*pocketful, questionnaire, isolate*, and *maniac*) indicated that the final forward diphone transitional probabilities for context items were mostly consistent with lexical context (3 out of the 4 items used). However, when we analyzed every context item used in previous studies, we discovered that diphone transitional probabilities could predict fewer than half of the 26 cases in which an LCfC effect was observed. This provides strong evidence that

transitional probabilities are not a good predictor of phoneme restoration in the context items and suggests that LCfC most likely results from feedback interactions between higher and lower levels of processing.

The current study highlights the need for thorough stimulus piloting prior to testing for LCfC effects. We find it particularly notable how few of the potential context-target pairs showed the expected compensation for coarticulation effect, with the effect only observed in 20 of the 130 pairs we piloted. While we observed robust phoneme restoration in a higher proportion of candidate contexts (10 of 35), this was also a lower rate that we would have expected. It may be that phoneme restoration and (non-lexical) compensation for coarticulation effects may be relatively difficult to observe, even though earlier papers did not report any difficulty in creating materials (Ganong, 1980; Mann & Repp, 1981; Repp & Mann, 1981, 1982). It is possible that the software we used to create continua (STRAIGHT; Kawahara et al., 2008) does not sufficiently isolate crucial acoustic details (a possibility we consider further below). We therefore suggest that future work could specifically investigate the acoustic details and other factors that may drive inconsistency in phoneme restoration and compensation for coarticulation effects.

As a starting point in this direction, we note that in the present study, compensation for coarticulation was only observed for target items that began with fricatives (specifically, /s/ or /ʃ/). Research by van der Zande, Jesse, and Cutler (2014) has suggested that there tends to be more variability in how talkers produce fricative sounds as compared, for instance, to stop consonants, which might explain why this particular set of phonemes was more sensitive to compensation for coarticulation effects. That being said, previous LCfC studies have successfully used non-fricative stimuli for targets (e.g., Elman & McClelland, 1988). One other possible reason we may have only observed compensation for coarticulation with fricative stimuli is because of the particular

morphing software (STRAIGHT; Kawahara et al., 2008) we used to create our stimuli; STRAIGHT is particularly well-suited to constructing continua between continuant speech sounds (e.g., fricatives, vowels) but may be less well-suited to making continua for obstruents (e.g., stop consonants), since the latter class of speech sounds is characterized by transient acoustic information that is difficult to model in STRAIGHT (McAuliffe, 2017).

It is further striking that while we began with a large set of potential stimuli (35 context continua and 32 target continua), only 4 context continua and 5 target continua survived piloting. As such, generalization may be a potential concern. However, we suggest instead that this speaks to the importance of extensive piloting for LCfC experiments. Critically, LCfC effects need not emerge in all stimuli, as this paradigm is a test case for the cognitive architecture of speech perception. Even if these results only emerge in a small number of stimuli, the phenomenon of LCfC is only predicted by an interactive model framework that allows for feedback. Any instance of a higher-level process directly influencing a low-level one necessarily satisfies the criteria of there being feedback present in a system. Feedforward models simply do not have the infrastructure to support this type of phenomena.

Though we observed robust LCfC effects in the current study – both in an initial sample and in a separate direct replication sample – larger LCfC effects could potentially constitute stronger evidence in favor of feedback, and we suggest that future work should consider manipulations that might increase the size of LCfC effects. For instance, given behavioral evidence that high-level context is particularly beneficial at intermediate signal-to-noise ratios (e.g., Davis, Ford, Kherif, & Johnsrude, 2011) as well as computational simulations showing that feedback is particularly beneficial for processing speech in noise (Magnuson et al., 2018), we might expect to find larger LCfC effects when speech is presented in noise. Alternatively, it might be informative

to test for LCfC in populations that may rely more heavily on top-down knowledge; for instance, previous work has suggested that older adults exhibit stronger effects of lexical knowledge on phonetic categorization (Mattys & Scharenborg, 2014; Pichora-Fuller, 2008; Rogers, Jacoby, & Sommers, 2012), and as such, it could be informative to test for LCfC in older adults specifically. Nevertheless, the present finding of robust LCfC (in young adults hearing speech in the absence of background noise) in and of itself provides strong evidence in favor of theories that include top-down feedback.

Our findings complement a growing body of results that suggest top-down effects in speech processing. For instance, recent behavioral work has shown that lexical status can influence whether incoming speech is perceived as a unified auditory stream or two segregated streams (Billig, Davis, Deeks, Monstrey, & Carlyon, 2013), consistent with earlier work demonstrating lexical influences on a listener's interpretation of ambiguous speech sounds (Ganong, 1980; Warren, 1970). One particularly compelling set of behavioral results comes from a set of studies by Samuel (1997, 2001), who, like Elman and McClelland (1988), tested for top-down effects by examining whether a lexically-restored phoneme could mediate a separate sub-lexical process. However, instead of leveraging compensation for coarticulation, Samuel leveraged the sub-lexical process of selective adaptation (Eimas & Corbit, 1973), a phenomenon in which exposure to a repeatedly presented stimulus (e.g., a clearly produced /d/) leads listeners to make fewer responses of that category on a subsequent test (e.g., fewer /d/ responses on a subsequent /d/-/t/ continuum). Samuel (1997) found that after repeated exposure to a lexically restored phoneme (e.g., the /d/ in *arma?illo*, where *?* indicates a phoneme replaced by white noise), listeners were less likely to report hearing that phoneme on a subsequent test continuum (e.g., fewer /d/ responses on a /b/-/d/ continuum); that is, a lexically restored phoneme had the same influence on the selective

adaptation process as an unambiguous phoneme did. In a second study, Samuel (2001) obtained similar results using an ambiguous phoneme that was a blend of two phonemes (e.g., a *s/sh* blend presented in the context *aboli#*) instead of using white noise. While Samuel argued that these results were best explained by interactive accounts, proponents of autonomous models have suggested that this may reflect a different sort of feedback (specifically, feedback for learning) that is distinct from the activation feedback at the core of interactive models (Norris, McQueen, & Cutler, 2003; but see Magnuson et al., 2018, for a counterargument on the basis of parsimony).

Neuroimaging data also support interactive models, as many researchers have found that when a listener hears an ambiguous speech sound, early neural activity is influenced by their interpretation of that sound; specifically, the neural response elicited by an ambiguous speech sound approximates the response elicited by an unambiguous variant of the perceived category (Bidelman, Moreno, & Alain, 2013; Getz & Toscano, 2019; Leonard, Baud, Sjerps, & Chang, 2016; Luthra, Correia, Kleinschmidt, Mesite, & Myers, 2020; Noe & Fischer-Baum, 2020). For instance, in a study by Noe and Fischer-Baum (2020), listeners were presented with word-nonword continua (e.g., *date-*tate, *dape-tape*) and were asked to categorize the initial sound as voiced (/d/) or voiceless (/t/). In a given block, listeners heard stimuli from only one continuum (e.g., *date-*tate*), and as such, each block was lexically biased toward one endpoint (e.g., /d/). The authors found that when listeners heard stimuli that were ambiguous between /d/ and /t/, the amplitude of the N100 (an early event-related potential) reflected the lexical bias of the block; that is, when listeners heard ambiguous stimuli in /d/-biased blocks, the N100 response resembled the response elicited by unambiguous /d/ sounds. Critically, the authors also demonstrated that this pattern of results could be modeled by the interactive TRACE model (McClelland & Elman, 1986) but not

by any strictly feedforward architectures. Overall, these results strongly favor theoretical accounts of spoken word recognition that allow for top-down feedback.

While the current investigation considered the domain of language processing in particular, our results may have implications for other domains, as debates over whether cognition involves top-down influences on perception recur in multiple domains (e.g., Firestone & Scholl, 2014, 2015). For instance, theories of visual object recognition have historically emphasized feedforward processing, but recent work suggests the importance of expanding theoretical accounts to include top-down processing as well (O'Callaghan, Kveraga, Shine, Adams, & Bar, 2017; Wyatte, Jilk, & O'Reilly, 2014). In our view, it is unlikely that feedback would exist in some domains but not others; rather, we suggest that top-down effects may be a core tenet of cognitive science, broadly speaking, consistent with the *interactive activation hypothesis* as articulated by McClelland, Mirman, Bolger, and Khaitan (2014).

Overall, a key goal of perception is to optimally integrate prior knowledge with sensory information (Clark, 2013; Davis & Sohoglu, 2020; Friston, 2010; Lupyan, 2015). By allowing for top-down feedback, interactive architectures can support an individual's ability to perform this integration, whether to infer the likely cause of a perceptual event (McClelland, 2013; McClelland et al., 2014) or to predict upcoming events (O'Callaghan et al., 2017; Panichello, Cheung, & Bar, 2013). Indeed, previous work in this domain illustrates that an appropriately parameterized interactive activation model can implement optimal Bayesian inference (McClelland et al., 2014). The present results provide strong evidence in favor of interactive theories, particularly in the domain of language processing, and provide key insights into the processes that underlie perception.

References

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388–407.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *NeuroImage*, *79*, 201–212.

Billig, A. J., Davis, M. H., Deeks, J. M., Monstrey, J., & Carlyon, R. P. (2013). Lexical influences on auditory streaming. *Current Biology*, *23*(16), 1585–1589.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

CMU Computer Science. (2020). CMU Pronouncing Dictionary. Retrieved August 25, 2020, from http://www.speech.cs.cmu.edu/cgi-bin/cmudict

Davis, M. H., Ford, M. A., Kherif, F., & Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through "top–down" processes? Evidence from time-resolved sparse fMRI. *Journal of Cognitive Neuroscience*, *23*(12), 3914–3932.

Davis, M. H., & Sohoglu, E. (2020). Three functions of prediction error for Bayesian inference in speech perception. In D. Poeppel, G. R. Mangun, & M. S. Gazzaniga (Eds.), *The Cognitive Neurosciences* (6th ed., pp. 177–189). Cambridge, MA: The MIT Press.

Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, *4*, 99–109.

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*(2), 143–165.

Firestone, C., & Scholl, B. J. (2014). "Top-down" effects where none should be found: The El Greco fallacy in perception research. *Psychological Science*, *25*(1), 38–46.

Firestone, C., & Scholl, B. J. (2015). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, (2016).

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110–125.

Getz, L. M., & Toscano, J. C. (2019). Electrophysiological evidence for top-down lexical influences on early speech perception. *Psychological Science*, *30*(6), 830–841.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363.

Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T., & Banno, H. (2008). Tandem-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. In *Proceedings*

of IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 3933–3936).

Kučera, H., & Francis, W. N. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, MA: Houghtin Mifflin.

Leonard, M. K., Baud, M. O., Sjerps, M. J., & Chang, E. F. (2016). Perceptual restoration of masked speech in human cortex. *Nature Communications*, *7*(1), 1–9.

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology*, *6*(4), 547–569.

Lupyan, G., Abdel Rahman, R., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences*.

Luthra, S., Correia, J. M., Kleinschmidt, D. F., Mesite, L. M., & Myers, E. B. (2020). Lexical information guides retuning of neural patterns in perceptual learning for speech. *Journal of Cognitive Neuroscience*, *32*(10), 2001–2012.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003a). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science*, *27*(2), 285–298.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003b). Lexical effects on compensation for coarticulation: A tale of two systems? *Cognitive Science*, *27*(5), 801–805.

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, 1–18.

Mann, V. A., & Repp, B. H. (1981). Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America*, *69*(2), 548–558.

Mattys, S. L., & Scharenborg, O. (2014). Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors. *Psychology and Aging*, *29*(1), 150–162.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

McAuliffe, M. (2017). STRAIGHT workshop. Retrieved September 28, 2020, from https://memcauliffe.com/straight_workshop

McClelland, J. L. (2013). Integrating probabilistic models of perception and interactive neural networks: A historical and tutorial review. *Frontiers in Psychology*, *4*, 1–25.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McClelland, J. L., Mirman, D., Bolger, D. J., & Khaitan, P. (2014). Interactive activation and mutual constraint satisfaction in perception and cognition. *Cognitive Science*, *38*(6), 1139–1189.

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), 1–18.

Montant, M. (2000). Feedback: A general mechanism in the brain. *Behavioral and Brain Sciences*, *23*(3), 340–341.

Noe, C., & Fischer-Baum, S. (2020). Early lexical influences on sublexical processing in speech perception: Evidence from electrophysiology. *Cognition*, *197*, 1–14.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*(3), 299–325.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238.

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience*, *31*(1), 4–18.

Norris, D., McQueen, J. M., & Cutler, A. (2018). Commentary on "Interaction in spoken word recognition models." *Frontiers in Psychology*, *9*, 1568, 1–3.

O'Callaghan, C., Kveraga, K., Shine, J. M., Adams, R. B., & Bar, M. (2017). Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Consciousness and Cognition*, *47*, 63–74.

Panichello, M. F., Cheung, O. S., & Bar, M. (2013). Predictive feedback and conscious visual experience. *Frontiers in Psychology*, *3*, 1–8.

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, *47*(SUPPL. 2).

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

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Repp, B. H., & Mann, V. A. (1981). Perceptual assessment of fricative-stop coarticulation. *Journal of the Acoustical Society of America*, *69*(4), 1154–1163.

Repp, B. H., & Mann, V. A. (1982). Fricative-stop coarticulation: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *71*(6), 1562–1567.

Rogers, C. S., Jacoby, L. L., & Sommers, M. S. (2012). Frequent false hearing by older adults: The role of age differences in metacognition. *Psychology and Aging*, *27*(1), 33–45.

Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology*, *32*(2), 97–127.

Samuel, A. G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science*, *12*(4), 348–351.

Samuel, A. G., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, *48*(2), 416–434.

Schnall, S. (2017a). No magic bullet in sight: A reply to Firestone and Scholl (2017) and Durgin (2017). *Perspectives on Psychological Science*, *12*(2), 347–349.

Schnall, S. (2017b). Social and contextual constraints on embodied perception. *Perspectives on Psychological Science*, *12*(2), 325–340.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2018). afex: Analysis of Factorial Experiments. R package version 0.21-2. https://CRAN.R-project.org/package=afex.

Stefanucci, J. K., & Geuss, M. N. (2009). Big people, little world: The body influences size perception. *Perception*, *38*(12), 1782–1795.

van der Zande, P., Jesse, A., & Cutler, A. (2014). Cross-speaker generalisation in two phoneme-level perceptual adaptation processes. *Journal of Phonetics*, *43*, 38–46.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*(3917), 392–393.

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, and Psychophysics*, *79*(7), 2064–2072.

Wyatte, D., Jilk, D. J., & O'Reilly, R. C. (2014). Early recurrent feedback facilitates visual object recognition under challenging conditions. *Frontiers in Psychology*, *5*, 1–10.

Appendix

Pitt and McQueen (1998) have suggested that positive results in LCfC studies may reflect influences of transitional probabilities rather than influences of the lexicon. To assess this claim, we computed frequency-weighted transitional probabilities for all context items used in previous LCfC studies (Elman & McClelland, 1988; Magnuson et al., 2003a; Pitt & McQueen, 1998; Samuel & Pitt, 2003) as well as the four items used in the current study (*questionnaire*, *pocketful*, *isolate* and *maniac*). Probabilities were computed using the SUBTLEX subtitle corpus (Brysbaert & New, 2009), which we cross-referenced with the Kučera and Francis (1982) database to reduce the word list to lemmas. Pronunciations not available in the database were filled in from the CMU Pronouncing Dictionary (CMU Computer Science, 2020).

Forward transitional probabilities (e.g., the probability that the next phoneme will be /s/ given the context *Christma_*) are provided in the tables below. The leftmost column indicates the pronunciation for each context (e.g., ɪsmʌ for *Christma_*). Some pronunciations may be specific to Providence or Pittsburgh dialects. Lexically consistent endings are shaded in gray. Transitional probabilities and lexical biases are correlated for approximately half the items (bolded black text), and transitional probabilities and lexical biases are in opposition for approximately half the items (bolded red text). An asterisk (*) is used to indicate the items (*juice*, *bush*) for which Pitt and McQueen (1998) did not observe lexically mediated compensation for coarticulation.

CHRISTMAS	s	ʃ	FOOLISH	s	ʃ
ʌ	**0.073068**	0.004041	ɪ	**0.103200**	0.030075
mʌ	**0.038507**	0.020243	lɪ	**0.128394**	0.056217
smʌ	**1.000000**	0.000000	ulɪ	0.061224	**0.387755**
ɪsmʌ	**1.000000**	0.000000	fulɪ	0.000000	**1.000000**
rɪsmʌ	**1.000000**	0.000000			
krɪsmʌ	**1.000000**	0.000000			

COPIOUS	s	ʃ	SPANISH	s	ʃ
ʌ	**0.073068**	0.004041	ɪ	**0.103200**	0.030075
iʌ	**0.253529**	0.003321	nɪ	**0.210175**	0.071802
piʌ	**0.027027**	0.000000	ænɪ	0.042480	**0.080367**
opiʌ	**0.105263**	0.000000	pænɪ	0.017857	**0.553571**
kopiʌ	**1.000000**	0.000000	spænɪ	0.000000	**1.000000**

RIDICULOUS	s	∫	ENGLISH	s	∫
ʌ	**0.073068**	0.004041	ɪ	**0.103200**	0.030075
lʌ	**0.083539**	0.000000	lɪ	**0.128394**	0.056217
ʊlʌ	**0.507576**	0.000000	glɪ	**0.119266**	0.500000
jʊlʌ	**0.566372**	0.000000	ŋglɪ	0.000000	**0.900826**
kjʊlʌ	**0.454545**	0.000000	ɪŋglɪ	0.000000	**1.000000**
ɪkjʊlʌ	**0.500000**	0.000000			
dɪkjʊlʌ	**1.000000**	0.000000			
ʌdɪkjʊlʌ	**1.000000**	0.000000			
rʌdɪkjʊlʌ	**1.000000**	0.000000			

ARTHRITIS	s	∫	ABOLISH	s	∫
ɪ	**0.103205**	0.030075	ɪ	**0.103200**	0.030075
tɪ	**0.099435**	0.022212	lɪ	**0.128394**	0.056217
aⁱtɪ	**0.068027**	0.000000	alɪ	**0.193160**	0.031666
raⁱtɪ	**0.055556**	0.000000	balɪ	0.000000	**0.204082**
θraⁱtɪ	**1.000000**	0.000000	ʌbalɪ	0.000000	**0.833333**
rθraⁱtɪ	**1.000000**	0.000000			
arθraⁱtɪ	**1.000000**	0.000000			

MALPRACTICE	s	∫	ESTABLISH	s	∫
ɪ	**0.103205**	0.030075	ɪ	**0.103205**	0.030075
tɪ	**0.099435**	0.022212	lɪ	**0.128394**	0.056217
ktɪ	**0.109726**	0.006234	blɪ	0.049765	**0.342723**
æktɪ	**0.219543**	0.010152	æblɪ	0.000000	**0.984556**
ræktɪ	**0.477901**	0.022099	tæblɪ	0.000000	**0.984556**
præktɪ	**0.545741**	0.025237	stæblɪ	0.000000	**1.000000**
lpræktɪ	**1.000000**	0.000000	ɪstæblɪ	0.000000	**1.000000**
ælpræktɪ	**1.000000**	0.000000			
mælpræktɪ	**1.000000**	0.000000			

CONTAGIOUS	s	ʃ		DISTINGUISH	s	ʃ
ʌ	**0.073068**	0.004041		ɪ	**0.103205**	0.030075
dʒʌ	**0.017606**	0.000000		wɪ	**0.066567**	0.092537
eⁱdʒʌ	**0.750000**	0.000000		gwɪ	0.136842	**0.312281**
teⁱdʒʌ	**1.000000**	0.000000		ŋgwɪ	0.136842	**0.312281**
nteⁱdʒʌ	**1.000000**	0.000000		ɪŋgwɪ	0.336207	**0.663793**
ʌnteⁱdʒʌ	**1.000000**	0.000000		tɪŋgwɪ	0.000000	**1.000000**
kʌnteⁱdʒʌ	**1.000000**	0.000000		stɪŋgwɪ	0.000000	**1.000000**
				ɪstɪŋgwɪ	0.000000	**1.000000**
				dɪstɪŋgwɪ	0.000000	**1.000000**

CONSENSUS	s	ʃ		EXTINGUISH	s	ʃ
ɪ	**0.103205**	0.030075		ɪ	**0.103205**	0.030075
sɪ	**0.215924**	0.000000		wɪ	0.066567	**0.092537**
nsɪ	**0.165083**	0.000000		gwɪ	0.136842	**0.312281**
ɛnsɪ	**0.053012**	0.000000		ŋgwɪ	0.136842	**0.312281**
sɛnsɪ	**0.159091**	0.000000		ɪŋgwɪ	0.336207	**0.663793**
nsɛnsɪ	**0.636364**	0.000000		tɪŋgwɪ	0.000000	**1.000000**
ʌnsɛnsɪ	**1.000000**	0.000000		stɪŋgwɪ	0.000000	**1.000000**
kʌnsɛnsɪ	**1.000000**	0.000000		kstɪŋgwɪ	0.000000	**1.000000**
				ɪkstɪŋgwɪ	0.000000	**1.000000**

PROMISE	s	ʃ		PUNISH	s	ʃ
ɪ	**0.103205**	0.030075		ɪ	**0.103205**	0.030075
mɪ	**0.130722**	0.073107		nɪ	**0.210175**	0.071802
amɪ	**0.236088**	0.000000		ʌnɪ	0.027778	**0.194444**
ramɪ	**0.808383**	0.000000		pʌnɪ	0.000000	**1.000000**
pramɪ	**1.000000**	0.000000				

KISS	s	ʃ		FISH	s	ʃ
ɪ	**0.103205**	0.030075		ɪ	**0.103205**	0.030075
kɪ	**0.064099**	0.005342		fɪ	0.071652	**0.115764**

MISS	s	ʃ		WISH	s	ʃ
ɪ	**0.103205**	0.030075		ɪ	**0.103205**	0.030075
mɪ	**0.130722**	0.073107		wɪ	0.066567	**0.092537**

BLISS	s	∫	BRUSH	s	∫
ɪ	**0.103205**	0.030075	ʌ	**0.092664**	0.013686
lɪ	**0.128394**	0.056217	rʌ	**0.078313**	0.075777
blɪ	0.049765	**0.342723**	brʌ	0.003759	**0.315789**

JUICE *	s	∫	BUSH *	s	∫
u	**0.045775**	0.035992	ʊ	0.000786	**0.015912**
dʒu	**0.023544**	0.000000	bʊ	0.000000	**0.070111**

POCKETFUL	l	r	QUESTIONNAIRE	l	r
ʊ	0.201159	**0.222473**	æ	0.036615	**0.100621**
fʊ	**0.462740**	0.000000	næ	**0.107984**	0.102240
tfʊ	**1.000000**	0.000000	ʌnæ	**0.643636**	0.210909
ɪtfʊ	**1.000000**	0.000000	tʃʌnæ	0.000000	**1.000000**
kɪtfʊ	**1.000000**	0.000000	stʃʌnæ	0.000000	**1.000000**
akɪtfʊ	**1.000000**	0.000000	ɛstʃʌnæ	0.000000	**1.000000**
pakɪtfʊ	**1.000000**	0.000000	wɛstʃʌnæ	0.000000	**1.000000**
			kwɛstʃʌnæ	0.000000	**1.000000**

ISOLATE	t	k	MANIAC	t	k
eⁱ	**0.220374**	0.119337	æ	0.044587	**0.110985**
leⁱ	**0.260887**	0.014317	iæ	0.044574	**0.416667**
ʌleⁱ	**0.429648**	0.000000	niæ	0.000000	**0.888889**
sʌleⁱ	**0.521008**	0.000000	eniæ	0.000000	**1.000000**
aⁱsʌleⁱ	**0.738462**	0.000000	meniæ	0.000000	**1.000000**