On the Relationship between Phoneme Activation and Lexically Mediated Compensation for Coarticulation*

Michael A. Shepherd

University of Southern California
mshepher@usc.edu

Elman and McClelland (1988) created tokens of 'Christmas' and 'foolish' with an identical *ambiguous* final fricative and presented them followed by an acoustic 'tapes'-'capes' continuum. Listeners gave more 'capes' responses following 'Christmas' than following 'foolish'—compensation for coarticulation (e.g. Mann and Repp 1981)—but, as the authors predicted using TRACE (McClelland and Elman 1986), the effect was weaker than that triggered by a lexically appropriate, unambiguous /s/ or /ʃ/. Magnuson et al. (2003) and Samuel and Pitt (2003) replicated this result, and Ohala and Feder (1994) and Samuel (1997) obtained significant but weaker effects using noise-replaced segments, which listeners are known to confidently perceive ("restore") in the words they report hearing (Warren 1970, Samuel 1981). According to Elman and McClelland's (1998) TRACE-based explanation, ambiguous and noise-replaced segments trigger weaker effects because they fail to activate a unique candidate phoneme—a situation which is later resolved by lexical activation—whereas acoustically cued, lexically appropriate segments provide both acoustic and lexical support to a single phoneme, which becomes more highly activated and triggers stronger effects. To test the assumptions underlying this explanation, we attempted to replicate and extend Magnuson et al.'s study, measuring the magnitude of coarticulatory compensation triggered by ambiguous, unambiguous, noise-replaced, and noise-added segments and eliciting a rating (on a seven-point scale from most /s/-like to most /ʃ/-like) of the effect-triggering phoneme on every trial. Unfortunately, we were unable to replicate Magnuson et al's (2003) findings, and our results were, therefore, largely inconclusive.

1. Introduction

Several studies—Elman and McClelland (1988), Ohala and Feder (1994), Samuel (1997), Magnuson et al. (2003), Samuel and Pitt (2003)—have shown that an acoustically ambiguous or noise-replaced (Warren 1970, Samuel 1981) segment whose identity is clear from the word in which it occurs can trigger effects such as compensation for coarticulation (e.g. Mann and Repp 1981), vowel normalization, and selective adaptation. These findings have been interpreted as supporting the assumption of McClelland and Elman's (1986) TRACE model that activation in speech perception proceeds not only forward—from the acoustic level to the phonemic level—but also backward—from the lexical level to the phonemic level.

Of interest to us is the fact that, in all of these studies, effects triggered by acoustically ambiguous or noise-replaced segments have been of a lesser magnitude than those triggered by unambiguous, acoustically intact segments. The reason for this, according to Elman and McClelland's (1988) TRACE-based account, is that ambiguous and noise-replaced segments fail to provide acoustic information supporting the activation of a unique candidate phoneme. Ambiguous segments activate multiple phonemes, e.g. both /s/ and /ʃ/ in Elman and McClelland (1988), Magnuson et al. (2003), and Samuel and Pitt (2003), while noise-replaced segments do not activate any. Later in the perception process, when an

* Special thanks to Justin Aronoff, Dani Byrd, Todd Haskell, Elsi Kaiser, Pat Keating, Jim Magnuson, Bob McMurray, and Colin Wilson. Errors are my own. Work supported by the USC Phonetics Laboratory.

ambiguous or noise-replaced segment's lexical context becomes activated, the lexically appropriate phoneme increases in activation by virtue of being part of a lexical item that has become activated. Crucially, however, such lexically cued phonemes never become as highly activated as their acoustically cued, lexically appropriate counterparts. This is because, in the latter case, both acoustic and lexical information converge on a single candidate phoneme. It is precisely this difference in the level of activation of the effect-triggering phoneme, Elman and McClelland (1988) claimed, that leads to the relative weakness of effects produced by ambiguous and noise-replaced segments.

Such an explanation makes a couple of assumptions that we are interested in testing. The first is that the magnitude of the effects in question is directly related to the level of activation of the effect-triggering phoneme. The second is that acoustically ambiguous and noise-replaced segments produce weaker effects than unambiguous, acoustically intact controls because they lack acoustic information supporting the activation of the crucial phoneme. Thus, while the effect-triggering phoneme does become activated in the ambiguous and noise-replaced conditions—as a result of top-down influence from the lexicon—its level of activation is never as high as that of a phoneme that receives both bottom-up and top-down support, as is the case in the unambiguous, acoustically intact conditions. Our experiment sought to test these assumptions.

We attempted to replicate and extend Magnuson et al. (2003), which itself was a replication of Elman and McClelland (1988). Both studies were built on a clever combination of two well-known perceptual effects—the Ganong Effect (Ganong 1980) and compensation for coarticulation (e.g. Mann and Repp 1981). We will begin with a review of the relevant literature, after which we will introduce our own experiment and discuss the results.

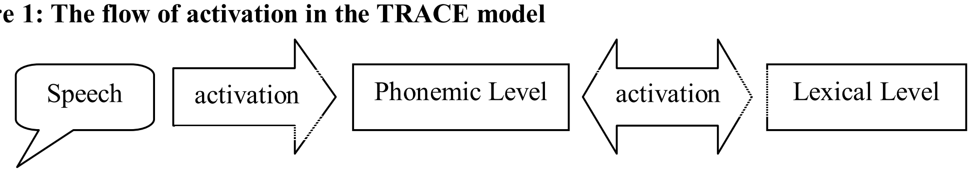
2. Literature review

The Ganong Effect (Ganong 1980) refers to the observation that listeners' phonemic category boundaries can be influenced by lexical context. Thus, when an acoustically ambiguous segment lies halfway between two categories—one of which, if selected, would form a word while the other would not—listeners will tend to perceive the ambiguous segment as belonging to the category that forms a word. For instance, when the elements of an acoustic [t]-[d] continuum are followed by [ɛst], listeners tend to label more of the continuum as belonging to the /t/ category, which forms the English word 'test', and make fewer /d/ responses, which would correspond to the non-word 'dest', relative to when they hear the continuum in isolation. Moreover, when the same continuum is paired with minimally different [ɛsk], the opposite pattern emerges—listeners make more /d/ categorizations, which correspond to the English word 'desk', and fewer /t/ responses, which make the non-word 'tesk'.

Compensation for coarticulation, as described by Mann and Repp (1981), refers to a (somewhat controversial) articulatory explanation for the observation that a segment which is ambiguous between /t/ and /k/ is more likely to perceived as /k/ following [s] than following [ʃ] or in isolation. Mann and Repp claimed that this occurs because listeners adjust their phonemic category boundaries so as to undo the articulatory influences of adjacent segments. Thus, since the place of articulation of a velar stop following /s/ is generally further forward on the palate than that of a velar stop in isolation or following a non-alveolar consonant such as /ʃ/, listeners automatically and unconsciously shift their alveolar-velar category boundary to compensate.

Elman and McClelland (1988) were interested in testing the predictions of their TRACE model of speech perception (McClelland and Elman 1986). Being an interactive model (as opposed to an autonomous or modular model), TRACE predicts not only that activation at the phonemic level leads to activation at the lexical level, but also the reverse, that is, that activation of a given lexical item leads to activation of its constituent phonemes. Figure 1 illustrates this two-way flow of activation between the phonemic and lexical levels.

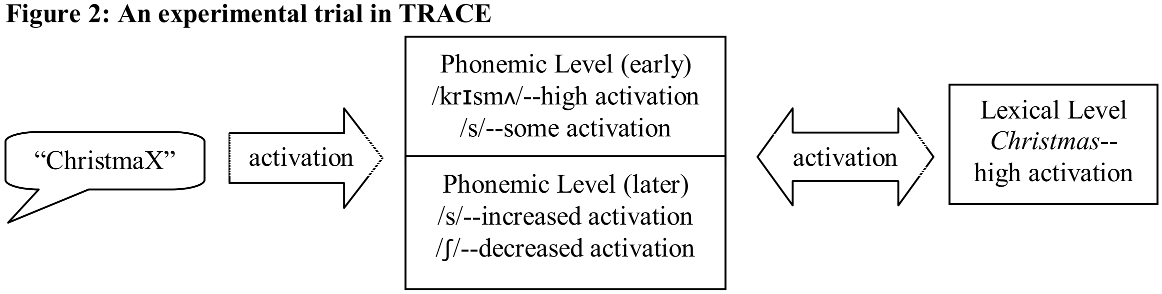
Figure 1: The flow of activation in the TRACE model



In order to test this prediction, Elman and McClelland designed an experiment which cleverly combined the Ganong Effect (Ganong 1980) and compensation for coarticulation (e.g. Mann and Repp 1981). They began by creating tokens of /s/- and /ʃ/-final English words, such as ‘Christmas’ and ‘foolish’, in which the last segment was acoustically halfway between canonical /s/ and /ʃ/. Despite being acoustically identical, listeners tended to categorize the ambiguous fricative as /s/ in the resulting tokens of /s/-final words and as /ʃ/ in the tokens of /ʃ/-final words, exemplifying the Ganong Effect. In a later experiment, the modified tokens were presented followed by items from an acoustic continuum ranging from a word with an initial alveolar stop, such as ‘tapes’, to an otherwise-identical word with an initial velar stop, such as ‘capes’. When subjects were asked to categorize the initial stop of the second word as either alveolar or velar, they gave more velar responses following the modified /s/-final words than the modified /ʃ/-final words. Elman and McClelland interpreted this as compensation for coarticulation. Moreover, since the modified words all ended in the same acoustically ambiguous fricative, the authors concluded that the differences in phoneme activation responsible for the different rates of alveolar and velar responses must have been due to top-down influences from the lexicon. Furthermore, as the authors also predicted using the TRACE model, the compensation for coarticulation effect obtained with the ambiguous fricative was of a lesser magnitude than that obtained with an unambiguous, lexically appropriate [s] or [ʃ]—an 11% shift in the mean percentage of /k/ responses following the ambiguous fricative as opposed to a 17% shift following the lexically appropriate, unambiguous fricatives.

Figure 2 provides an (admittedly simplistic) illustration of what a trial from Elman and McClelland’s experiment might look like in TRACE. First, some acoustic input (in this case “Christmas” with an ambiguous final fricative) is perceived. Activation then proceeds to the phonemic level—labeled “(early)” in figure 2—and the unambiguously cued phonemes (all but the final /s/) reach a high level of activation. As for the final phoneme, both /s/ and /ʃ/ reach roughly equal, intermediate levels of activation as candidates for the phoneme cued by the ambiguous fricative. This pattern of activation passes to the lexical level, where it leads to the selection of the lexical item *Christmas*, which becomes highly activated. The activation of *Christmas* then feeds back to the phonemic level—labeled “(later)” in figure 2—leading to increased activation of /s/ relative to /ʃ/ as a candidate for the phoneme cued by the ambiguous fricative. Since both /s/ and /ʃ/ are competing for recognition, the increase in activation of /s/ inhibits /ʃ/, leading to decreased activation of /ʃ/ relative to /s/. As a result of the higher level of activation of /s/ than /ʃ/, the perceptual system adjusts the alveolar-velar boundary (compensates for coarticulation). It therefore becomes more likely that a stop which is ambiguous between /t/ and /k/ will be perceived as /k/.

Figure 2: An experimental trial in TRACE



Subsequent studies have used the *phonemic restoration* effect (Warren 1970, Samuel 1981 et seq.), in which listeners confidently perceive (“restore”) noise-replaced segments in the words they report hearing. These studies have found that vowel normalization and selective adaptation effects can be triggered by such “restored” segments but are likewise weaker than those produced by acoustically intact controls. Ohala and Feder (1994) found that listeners normalize vowel quality in the context of noise-replaced consonants much as they do in the context of acoustically intact ones. However, the shift in the identification function in the noise-replaced condition was roughly half of that observed in the acoustically intact condition. Likewise, Samuel (1997) obtained selective adaptation effects with noise-replaced consonants. Again, however, the effects were about half as large as those produced by acoustically intact consonants—an 8.1% shift in response category in the noise-replaced condition versus a 16% shift in the intact condition.

More recently, Magnuson et al. (2003) replicated Elman and McClelland’s (1988) study in order to address concerns about the original result. Pitt and McQueen (1998) had suggested that the apparent compensation for coarticulation may have been due to biases in the diphone probabilities of the final VCs of the /s/- and /ʃ/-final stimulus words, rather than, as Elman and McClelland had claimed, genuine top-down influences on phoneme activation. This was based on Pitt and McQueen's (mistaken) observation that /s/ occurs more frequently than /ʃ/ following the final vowel in 'Christmas' and that the opposite is true following the final vowel in 'foolish'.

In response to Pitt and McQueen's concerns, Magnuson et al. recalculated the diphone probabilities of the final VCs of Elman and McClelland's /s/- and /ʃ/-final stimulus words. They found that Elman and McClelland's stimuli were, in fact, biased *against* the results they obtained. Despite this, and despite the fact that Elman and McClelland had also reported an experiment in which these VC combinations were reversed, Magnuson et al. went on to replicate Elman and McClelland's study using the words 'bliss' and 'brush', whose vowels favor a following /ʃ/ and /s/, respectively.

As in Elman and McClelland's study, the final fricatives of 'bliss' and 'brush' were replaced with acoustically ambiguous ones and presented followed by items from an acoustic 'tapes'-'capes' continuum. As before, compensation for coarticulation was obtained with the ambiguous fricatives but was weaker than that brought about by a lexically appropriate, unambiguous /s/ or /ʃ/—a 6% shift in the mean percentage of /k/ responses in each of the two ambiguous conditions versus a 20% shift following lexically appropriate, unambiguous fricatives.

Samuel and Pitt (2003) also replicated Elman and McClelland's (1988) results. Of particular interest to us is the fact that, as in all such studies we have seen, when compensation for coarticulation was triggered by an acoustically ambiguous, lexically cued fricative, it was of a lesser magnitude than when it was triggered by an acoustically intact, lexically appropriate segment. Samuel and Pitt also incorporated another condition of interest, namely, one using an unambiguous but lexically inappropriate fricative. Compensation for coarticulation also occurred in this condition, but with similarly reduced magnitude relative to that obtained with the unambiguous, lexically appropriate fricatives. These results are consistent with Elman and McClelland's (1988) TRACE-based account insofar as the convergence of acoustic and lexical support led to the highest level of activation of the compensation-triggering phoneme (as evidenced by the most compensation for coarticulation), whereas when activation of the compensation-triggering phoneme received only acoustic support or only lexical support, the magnitude of the effect was dramatically reduced.

3. Method

Our experiment was intended to both replicate and extend Magnuson et al. (2003) by measuring the magnitude of compensation for coarticulation in the original unambiguous and ambiguous conditions as well as in novel noise-replaced and noise-added conditions. We also elicited more detailed information, replacing binary /s/ or /ʃ/ categorization of the final fricative of the first word ('bliss' or 'brush') with a seven-point rating scale ranging from most /s/-like to most /ʃ/-like. This rating was elicited on every trial immediately prior to the /t/-/k/ categorization. This was to allow us to test the first assumption of Elman and McClelland's (1988) TRACE-based explanation for why acoustically ambiguous and noise-replaced segments produce weaker effects than unambiguous, acoustically intact controls, namely, that the magnitude of effects is directly related to the level of activation of the effect-triggering phoneme. We planned to use the subjects' seven-point ratings of the fricatives to test for a correlation between the magnitude of compensation for coarticulation (the ratio of /t/ responses to /k/ responses) and the level of activation of /s/ relative to /ʃ/ (as reflected by subjects' ratings of the effect-triggering fricative).[1]

The noise-added and noise-replaced conditions were intended to provide a means of testing the second assumption of Elman and McClelland's (1988) explanation, namely, that ambiguous and noise-replaced segments produce weaker effects than unambiguous, acoustically intact controls because they lack acoustic information supporting the activation of the effect-triggering phoneme. It was thought that the acoustic information present in the noise-added condition would provide bottom-up support for the lexically

[1] Samuel (1996) used subjects' intactness ratings (on an eight-point scale) of noise-added and noise-replaced stimuli relative to acoustically intact stimuli as a measure of the discriminability of the noise-added stimuli vis-à-vis their noise-replaced counterparts, and regarded discriminability as an inverse measure of phonemic restoration and activation.

appropriate fricative. This would in turn lead to greater activation and, thus, stronger compensation for coarticulation than in the noise-replaced condition, in which the crucial phoneme would receive only top-down support.[2]

3.1. Materials

Jim Magnuson and Bob McMurray (of Magnuson et al. 2003) kindly agreed to share their most recent stimuli with us. While these were not the same stimuli they had used in the original study, Magnuson and McMurray (personal communication) said they had successfully obtained lexically mediated compensation for coarticulation with them and, despite some "bugs," said they considered the new materials much more reliable and were much more confident in them.

According to Bob McMurray (personal communication), the original stimuli were constructed using natural speech and sample averaging. This means that a fricative which was ambiguous between /s/ and /ʃ/ consisted of the acoustics of both [s] and [ʃ] at half amplitude—a sound which could not be produced by an actual vocal tract. The stimuli which Magnuson and McMurray shared (and which we used) were new synthetic materials constructed by McMurray using the KlattWorks (McMurray, in preparation) interface to the Klatt (1980) synthesizer. In these stimuli, each of the ambiguous members of the /s/-/ʃ/ continuum was a single, acoustically intermediate segment, rather than a combination of the two unambiguous endpoints at particular amplitudes. Despite these differences, it is worth noting that the initial CCVs of 'bliss' and 'brush' were still produced in isolation and later combined with the fricatives so as to avoid any coarticulatory cues in the vowels. Thus, the unambiguous tokens of 'bliss' and 'brush' were created by combining /blɪ/ with an unambiguous /s/ and /brʌ/ with an unambiguous /ʃ/. Likewise, the ambiguous tokens were created by combining both /blɪ/ and /brʌ/ with ambiguous fricatives.

The materials we received from Magnuson and McMurray consisted of a total of 70 WAV files—all possible combinations of the two lexical contexts (/blɪ/ and /brʌ/) combined with fricatives from a five-step /s/-/ʃ/ continuum and followed by items from a seven-step 'tapes'-'capes' continuum. Based on the results of several pilot experiments, and consistent with Elman and McClelland (1988) and Magnuson et al. (2003), we elected to use unambiguous /s/ and /ʃ/ only where lexically appropriate, i.e. to form 'bliss' and 'brush' but never 'blish' and 'bruss'.[3] Pilot results also showed that, of the three intermediate steps of the /s/-/ʃ/ continuum, the one closest to the /s/ end was most ambiguous, so the other two were not used.

To create the noise-replaced stimuli, we used Audacity to delete the fricative portions of the nonce-word tokens ('blish' and 'bruss') and replace them with white noise comparable in duration and average amplitude. This was done using the nonce-word tokens as a safeguard, such that if any portion of the fricative inadvertently remained, it would work against the lexical restoration effect. We created the noise-added stimuli by taking the same duration- and amplitude-matched white noise that was used to create the noise-replaced tokens and digitally adding it to the fricative portion of each of the unambiguous and ambiguous tokens selected for use in the experiment.

As in Magnuson et al.'s (2003) original experiment, each of the various tokens of 'bliss' and 'brush' was followed by every step of an acoustic continuum ranging from an alveolar-initial word to an otherwise-identical velar-initial word. However, whereas they used a 'tapes'-'capes' continuum, we used Audacity to

---

[2] This assumption is not without controversy. Warren (1984) argued that phonemic restoration derives from a process of auditory induction, in which the excised speech sound is not restored, per se, but rather perceived within the noise such that the portion of the noise that matches the frequency of the excised speech is perceived as being the speech sound, and the higher and lower frequency portions of the noise as perceived as background noise. If phonemic restoration is indeed a case of auditory induction, then the noise-replaced stimuli do provide bottom up support for both /s/ and /ʃ/, much as the ambiguous stimuli are assumed to do. Repp (1992), on the other hand, argued that phonemic restoration is an auditory illusion accompanying completion of an abstract phonological template. If this is the case, no relevant acoustic information is derived from the noise, and our assumption stands. Samuel (1996), however, suggests that both auditory induction and illusion play a role in phonemic restoration, but that when fricatives and stops are replaced with white noise, bottom-up factors (induction) dominates.

[3] When 'blish' and 'bruss' were left out, not only was the overall Ganong Effect stronger, but subjects' ratings of the noise-added and noise-replaced fricatives were more similar to each other. Following Samuel (1981), we took this increased similarity in rating as evidence of a stronger lexical restoration effect insofar as it indicates subjects' inability to distinguish actual fricatives in noise from perceptually restored ones.

delete the final /s/ from ours (resulting in a ‘tape’-‘cape’ continuum) because we were concerned about the potential effects of subjects hearing a clear /s/ on every trial.[4]

Lastly, in order to reduce the total number of stimulus items, we elected to use only the middle five steps of the seven-point ‘tape’-‘cape’ continuum and omit the two endpoints. In pilot tests, only the middle three steps were ever identified as both ‘tape’ and ‘cape’ (the first two steps nearest to the ‘tape’ end of the continuum were always labeled ‘tape’ and the two nearest to the ‘cape’ end were always labeled ‘cape’), so this still left endpoints that we considered sufficiently unambiguous.[5]

Thus, in the final stimulus set, both /blɪ/ and /brʌ/ were combined with the ambiguous (between /s/ and /ʃ/) fricative, the ambiguous fricative with noise added to it, and the same noise with no speech sound underlying it. However, only /blɪ/ was followed by the unambiguous /s/ (both with and without noise added to it), and only /brʌ/ was followed by the unambiguous /ʃ/ (both with and without noise added to it), such that there were no unambiguous nonce words. All ten of these were paired with the middle five steps of the ‘tape’-‘cape’ continuum, resulting in a total of 50 two-word stimulus items. The following table summarizes this.[6]

Table 1: Experimental stimuli

<table>
<tr><td>only /blɪ/</td><td>unambiguous /s/,
unambiguous /s/ + noise</td><td rowspan="3">all combined with all five steps
of the ‘tape’-‘cape’ continuum</td></tr>
<tr><td>both /blɪ/
and /brʌ/</td><td>ambiguous /s/-/ʃ/,
ambiguous /s/-/ʃ/ + noise,
noise only</td></tr>
<tr><td>only /brʌ/</td><td>unambiguous /ʃ/,
unambiguous /ʃ/ + noise</td></tr>
</table>

3.2. Procedure

The experiment was run using ePrime 1.1 (Psychology Software Tools, Inc). On each trial, participants heard one of the 50 two-word stimulus items described above and were required to make two responses. First, they pressed one of seven specially labeled keys on the keyboard to rate the final fricative of the first word on a seven-point scale from most /s/-like to most /ʃ/-like. Subjects then pressed one of two other specially labeled keys to indicate whether the second word was ‘tape’ or ‘cape’. The first response was made using the lower row of keys—from Z to M—with the left end of the scale—Z—indicating most /s/-like and the right end—M—most /ʃ/-like. The second response was made using the D and H keys, which are in the middle row of the keyboard, centered above the second and third, and fifth and sixth, steps of the /s/-/ʃ/ rating scale, respectively.

On each trial, one of the 50 two-word stimulus items was played over headphones. Immediately afterwards, a large yellow “+” sign appeared in the middle of the black screen to remind subjects that they must first rate the final fricative of the first word on the /s/-/ʃ/ scale (the keys used for this task were labeled with yellow squares of paper). As soon as they made a response, the yellow “+” sign changed to blue, reminding subjects that they must then categorize the second word as either ‘tape’-‘cape’ (the keys used for this second task were labeled with blue squares of paper). After they made the second response, the screen went black. Following the 1000ms inter-stimulus interval, another sound item was played and the cycle

[4] Pilot tests confirmed that using a ‘tapes’-‘capes’ continuum resulted in the ambiguous final fricative of ‘bliss’ and ‘brush’ being judged more /ʃ/-like than with a ‘tape’-‘cape’ continuum—a selective adaptation effect. We also confirmed with Bob McMurray (personal communication) that he could see no problem with using a ‘tape’-‘cape’ continuum rather than one ranging from ‘tapes’ to ‘capes’.

[5] Note that while other studies (e.g. Samuel 1997, Magnuson et al. 2003) used full continua, the authors limited their analyses to the middle seven of nine steps and the middle four of eight steps, respectively.

[6] In the experiment, the twenty non-noise tokens (both ambiguous and unambiguous) were presented twice per block, making a total of 70 items per block—30 with noise and 40 without. This was done so as to avoid a situation in which stimuli with noise would outnumber stimuli without noise.

repeated itself.[7] There was no need for subjects to look at the screen during the experiment, and many did not. However, pilot subjects using a version of the experiment in which the screen remained black throughout suggested that it would be comforting to have some sort of visual confirmation that responses were actually being recorded. There was no time limit to respond, but most subjects made both responses within a second or two. Reaction times were not recorded.

The experiment began with 40 practice trials.[8] In order to ease subjects into the noise and ambiguity, the first four practice trials always consisted of the four combinations of unambiguous ‘bliss’ or ‘brush’ followed by either the ‘tape’ or ‘cape’ endpoint, in random order. These were followed by 36 additional practice trials randomly selected from the same 70-item list used for the experimental blocks. Following the practice trials, there were four experimental blocks. In each block, all 70 items from the trial list were presented in random order, resulting in a total of 280 experimental trials per subject.

3.3. Participants

Twenty adult native English speakers with self-reported normal hearing from the USC undergraduate population were paid for their participation.

4. Predictions

Before presenting the results of our experiment, let us briefly review our predictions. First, lexical context (/blɪ/ or /brʌ/) was expected to have a significant effect on subjects’ seven-point, /s/-/ʃ/ rating of the ambiguous fricative, the ambiguous fricative with noise added to it, and the noise with no underlying speech sound. Specifically, subjects were expected to give more /s/-like ratings in the context of /blɪ/ and more /ʃ/-like ratings in the context of /brʌ/—exemplifying the Ganong Effect (Ganong 1980). Our second prediction was that subjects would give significantly more ‘cape’ responses following unambiguous ‘bliss’ than unambiguous ‘brush’—exemplifying compensation for coarticulation (e.g. Mann and Repp 1981).

In the ambiguous, non-noise condition, subjects were expected to give significantly more ‘cape’ responses in the /blɪ/ context than in the /brʌ/ context. However, as in all previous studies measuring the magnitude of such effects triggered by acoustically ambiguous and noise-replaced segments, the difference in the mean percentage of ‘cape’ responses was expected to be smaller than in the baseline condition. These two results would constitute a replication of Magnuson et al. (2003), and would support the claim that lexical activation can exert a top-down influence on phoneme activation and lead to greater activation of the lexically appropriate fricative despite ambiguous acoustic cues. Subjects were also expected to give less endpoint-like fricative ratings in the ambiguous condition (again, relative to the unambiguous baseline condition), indicating that the effect-triggering fricative was less highly activated in the ambiguous condition than in the unambiguous condition.

In the three conditions with noise—unambiguous, ambiguous, and noise-only—subjects were expected to experience the phonemic restoration effect (Warren 1970, Samuel 1981) and, thus, report hearing an acoustically intact, lexically appropriate fricative regardless of whether one was actually present in the acoustic signal. It was further expected that, independent of subjects’ conscious perceptions of the stimuli, the [s] + noise and [ʃ] + noise conditions would provide more acoustic information supporting bottom-up activation of /s/ and /ʃ/, respectively, than either the ambiguous fricative + noise or noise-only conditions because an unambiguous, lexically appropriate fricative would be present in the noise. If the unambiguous fricative + noise conditions did lead to greater activation of the effect-triggering fricative, we would expect to see stronger compensation for coarticulation in these conditions than in the ambiguous fricative + noise and noise only conditions.

---

[7] A 1000ms inter-stimulus interval was selected based on feedback from pilot subjects who experienced versions of the experiment with both longer and shorter intervals. Magnuson et al. (2003) used an inter-stimulus interval of 800ms.

[8] This was based on the average number of trials needed for pilot subjects to begin making full and consistent use of the seven-point /s/-/ʃ/ rating scale. Magnuson et al. (2003) gave their subjects 36 practice trials.

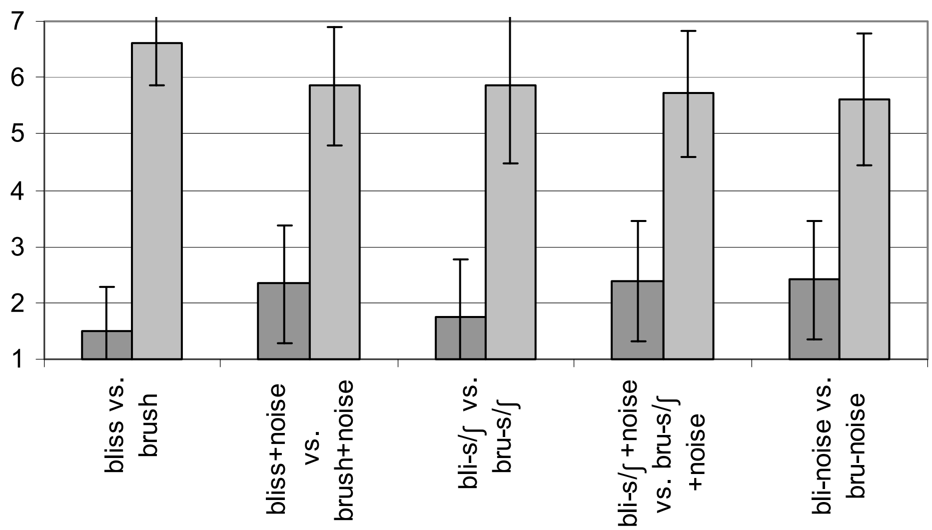
5. Results and analysis

5.1. The Ganong effect

We began by examining subjects' seven-point, /s/-/ʃ/ fricative ratings in order to determine whether lexical context (/blɪ/ or /brʌ/) influenced subjects' perception of the ambiguous fricative, the ambiguous fricative with noise added, and the same noise with no speech sound underlying it. Looking first at the unambiguous conditions, the final /s/ in unambiguous 'bliss' (labeled "bliss" in the graph below) received a mean rating of 1.5 (SD = 0.8) on the seven-point /s/-/ʃ/ scale (one being most /s/-like), while the final /ʃ/ in unambiguous 'brush' (labeled "brush" in the graph below) received a mean rating of 6.6 (SD = 0.8). With noise added, the final /s/ in unambiguous 'bliss' ("bliss+noise") received a mean rating of 2.3 (SD = 1.0), while the final /ʃ/ in unambiguous 'brush' ("brush+noise") received a mean rating of 5.9 (SD = 1.1).

As predicted, lexical context strongly affected responses to the ambiguous fricative, the ambiguous fricative with noise-added, and the noise with no speech sound underlying it. The ambiguous fricative in the context of /blɪ/ ("bli-s/ʃ" in the graph below) received a mean rating of 1.7 (SD = 1.1), while the same fricative in the context of /brʌ/ ("bru-s/ʃ") received a mean rating of 5.9 (SD = 1.4). With noise added, the ambiguous fricative in the context of /blɪ/ ("bli-s/ʃ+noise") received a mean rating of 2.4 (SD = 1.1), while the same fricative in the context of /brʌ/ ("bru-s/ʃ+noise") received a mean rating of 5.7 (SD = 1.1). Lastly, the noise-replaced fricative in the context of /blɪ/ ("bli-noise") received a mean rating of 2.4 (SD = 1.0), while the same noise in the context of /brʌ/ ("bru-noise") received a mean rating of 5.6 (SD = 1.2). These ratings are summarized in figure 3.

Figure 3: Mean /s/-/ʃ/ ratings by condition



In order to verify that the lexical context effect held across participants and 'tape'-'cape' continuum steps, we ran three two-factor, repeated-measures ANOVAs using StatView (one on the ambiguous non-noise condition, one on the ambiguous noise-added condition, and one on the noise only condition) with 2 × 5 levels (lexical context × 'tape'-'cape' continuum step) on the /s/-/ʃ/ ratings. The main effect of context was significant in all three conditions (ambiguous non-noise $F(1,19) = 244.2$, $p < .0001$; ambiguous noise-added $F(1,19) = 114.9$, $p < .0001$; noise only $F(1,19) = 105.5$, $p < .0001$) indicating that lexical context (/blɪ/ or /brʌ/) had a significant effect on subjects' /s/-/ʃ/ ratings of the ambiguous fricative, the ambiguous fricative with noise-added, and noise with no underlying speech sound—exemplifying the Ganong Effect (Ganong 1980). The effect of 'tape'-'cape' continuum step was not significant in any of the three conditions (ambiguous, non-noise $F(4,76) = 1.7$; ambiguous, noise-added $F(4,76) = 2.0$; noise only $F(4,76) = 1.6$) nor was the interaction of context and continuum step (ambiguous non-noise $F(4,76) < 1$; ambiguous noise-added $F(4,76) = 1.9$; noise only $F(4,76) < 1$), indicating that the effect of lexical context on fricative rating was the same regardless of which item (from the 'tape'-'cape' continuum) followed.

5.2. Compensation for coarticulation

Having looked at subjects' fricative ratings, subjects' 'tape'-'cape' categorizations were tested for evidence of compensation for coarticulation. Figures 4 through 8 below show the effect of lexical context on the subjects' 'tape'-'cape' responses at each step of the 'tape'-'cape' continuum. The two curves represent the two lexical contexts (/blɪ/ and /brʌ/), with 'tape'-'cape' continuum step on the x-axis and percentage of 'cape' responses on the y-axis. We ran five two-factor, repeated-measures ANOVAs (one for each condition) with 2 × 5 levels (lexical context × 'tape'-'cape' continuum step) on the percentage of 'cape' responses. The results for each of the five conditions which will be discussed in turn.

Figure 4 shows the percentage of 'cape' responses in the baseline condition, i.e. following unambiguous 'bliss' and unambiguous 'brush'. As predicted, the 'bliss' curve is higher than the 'brush' curve at every step of the continuum, indicating that subjects made more 'cape' responses following 'bliss' than following 'brush'—exemplifying compensation for coarticulation (e.g. Mann and Repp 1981). The main effect of context on the rate of 'cape' responses was significant ('bliss' = 46% 'cape' responses, 'brush' = 35% 'cape' responses; $F(1,19) = 10.3$, $p < .005$).[9] Not surprisingly, the main effect of 'tape'-'cape' continuum step (which ranged from unambiguous 'tape' to unambiguous 'cape') on the rate of 'cape' responses was also significant ($F(4,76) = 131.1$, $p < .0001$). There was also a significant interaction between context and continuum step ($F(4,76) = 4.4$, $p < .005$). This is not surprising since, as figure 4 indicates, the effect of context on 'tape'-'cape' responses was greatest in the ambiguous middle portion of the 'tape'-'cape' continuum.

Figure 4: Unambiguous (baseline)

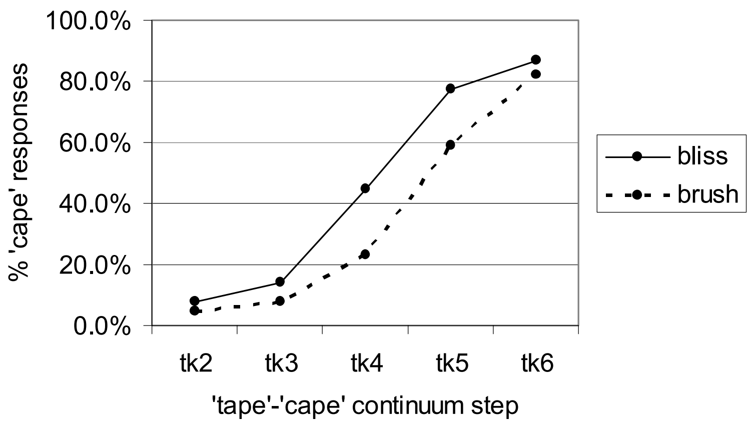


Figure 5 shows the percentage of 'cape' responses in the unambiguous, noise-added condition, which is identical to the baseline condition except that white noise was added to the fricatives in 'bliss' and 'brush'. Unlike in the graph of the baseline condition, the 'bliss' curve is actually slightly lower than the 'brush' curve at all but the first step of the 'tape'-'cape' continuum, indicating that subjects made marginally more 'tape' responses following 'bliss' than following 'brush'—the opposite of the predicted result—but this main effect of context was not significant ($F(1,19) = 1.9$). As in the baseline condition, there was a significant main effect of continuum step ($F(4,76) = 45.9$, $p < .0001$). There was no interaction between continuum step and context ($F(4,76) = 1.2$).

[9] In the unambiguous and unambiguous, noise-added conditions, lexical context (/bli/ or /bru/) was confounded with the acoustics of the fricative since unambiguous /s/ only followed /bli/ and unambiguous /ʃ/ only followed /bru/.

Figure 5: Unambiguous with noise added

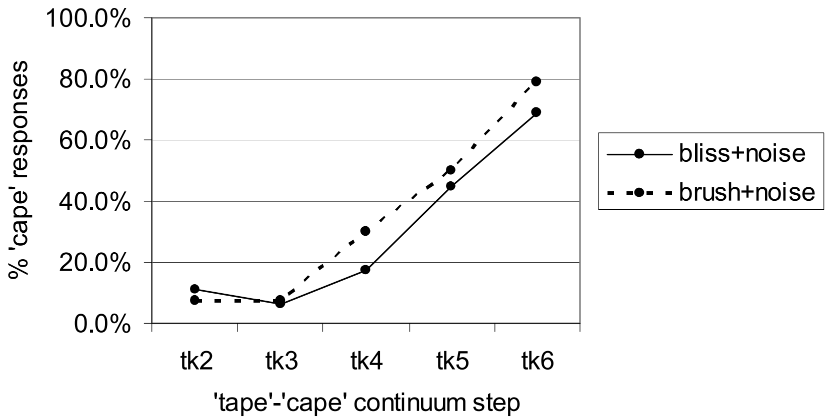


Figure 6 shows the percentage of ‘cape’ responses in the ambiguous, non-noise condition, which was intended to replicate Magnuson et al.’s (2003) finding of contextually mediated compensation for coarticulation with an ambiguous fricative. As the graph shows, however, there was no consistent relationship between lexical context and the rate of ‘cape’ responses across ‘tape’-‘cape’ continuum steps. Indeed, the main effect of lexical context was not significant ($F(1,19) < 1$). As before, there was a significant main effect of continuum step ($F(4,76) = 105.3$, $p < .0001$), and no interaction between continuum step and context ($F(4,76) = 2.1$).

Figure 6: Ambiguous non-noise

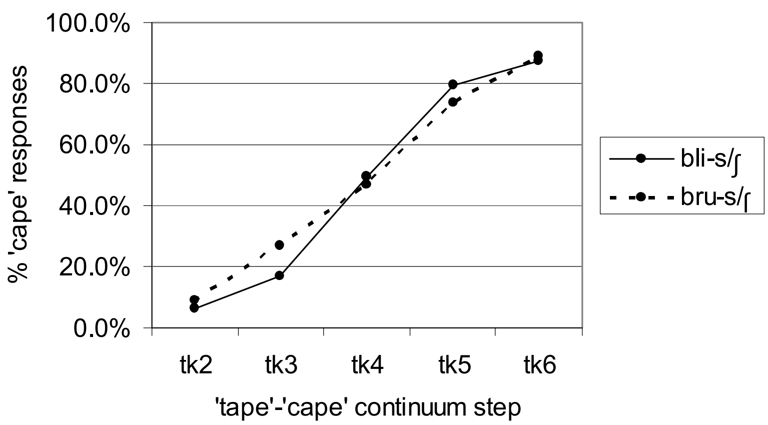
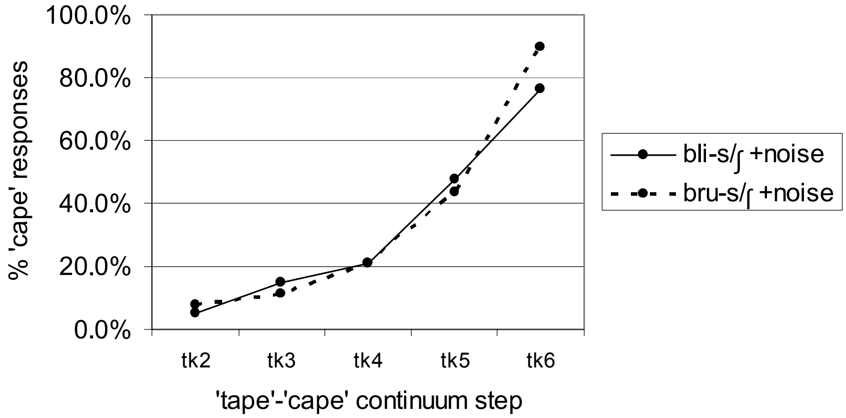


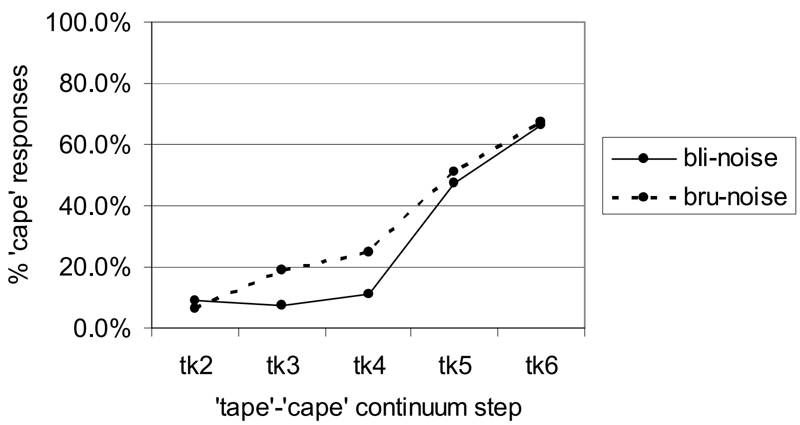
Figure 7 shows the percentage of ‘cape’ responses in the ambiguous, noise-added condition. As was the case in the ambiguous, non-noise condition, there was no consistent relationship between lexical context and the rate of ‘cape’ responses across ‘tape’-‘cape’ continuum steps, such that the main effect of context was not significant ($F(1,19) < 1$). There was, of course, a significant main effect of continuum step ($F(4,76) = 45.7$, $p < .0001$). There was no interaction between continuum step and context ($F(4,76) = 1.7$).

Figure 7: Ambiguous with Noise Added



Lastly, figure 8 shows the percentage of ‘cape’ responses in the noise-only condition. As was the case in the unambiguous, noise-added condition, the ‘bliss’ curve is lower than the ‘brush’ curve at all but the first step of the ‘tape’-‘cape’ continuum, indicating that subjects made more ‘tape’ responses following ‘bliss’ than following ‘brush’. However, this main effect of context was not significant ($F(1,19) = 1.9$). As in the four other conditions, the main effect of continuum step was significant ($F(4,76) = 36.1$, $p < .0001$). There was no interaction between continuum step and context ($F(4,76) = 1.4$).

Figure 8: Noise only



6. Discussion and conclusion

Despite what looked to be a very promising Ganong Effect and good compensation for coarticulation in the baseline condition, our experiment failed to replicate, much less extend, Magnuson et al.’s (2003) finding of compensation for coarticulation mediated by lexical context. Why might this have occurred? It is possible that task complexity was at issue. Unlike Magnuson et al. (2003), in which subjects made a categorical /s/ or /ʃ/ response, our subjects were asked to rate the quality of the fricative on a seven-point scale from /s/ to /ʃ/. Time-course factors may also have been involved insofar as subjects rated the fricative first and then categorized the stop. A third possibility is that omitting the two endpoints of the ‘tape’-‘cape’ continuum and using only the potentially ambiguous items made the ‘tape’-‘cape’ categorization task too difficult.

What implications, if any, do our results have for the two assumptions of the TRACE model that we hoped to test? The first assumption was that the magnitude of effects is directly related to the level of activation of the effect-triggering phoneme. Given the huge discrepancy between subjects’ ratings of the fricatives, which showed a strong effect of lexical context, and their categorizations of the following words, which failed to show the expected pattern of compensation for coarticulation, two possibilities exist. One is that Elman and McClelland were wrong, and the magnitude of effects is not directly related to the level of activation of the effect-triggering phoneme. It seems more likely, however, that it was incorrect to assume that an offline measure such as subjects' conscious fricative ratings would reflect the level of activation of /s/ relative to /ʃ/ (or perhaps that other factors, such as a post-perceptual bias towards real words or simply task complexity, obscured these ratings). In either case, it seems clear that in order to test the TRACE model's assumption about the relationship between phoneme activation and effects, a more direct method of measuring phoneme activation (perhaps an online measure, such as eye-tracking or brain imaging, quite possibly incorporated into an entirely different task) would be needed.

The second assumption of TRACE that we were interested in testing is that acoustically ambiguous and noise-replaced segments produce weaker effects than unambiguous, acoustically intact controls because they lack acoustic information supporting greater activation of the effect-triggering phoneme relative to its competitors. We had hoped that the noise-added and noise-replaced conditions would provide a crucial test of this assumption. The acoustic information present in the noise-added condition (and crucially absent from the noise-replaced condition) was intended to provide bottom-up support for the lexically appropriate fricative. According to TRACE, this bottom-up support should lead to greater activation of the fricative and, thus, stronger compensation for coarticulation than in the noise-replaced condition, in which the crucial phoneme would receive only top-down support. However, since our results failed to show significant compensation for coarticulation in either of these two conditions, were are unable to compare the relative magnitude of the effect in each, and therefore have nothing to say with regard to the second assumption.

Thus, while our experiment successfully replicated the Ganong Effect and compensation for coarticulation, it failed to replicate or extend the results of Magnuson et al.'s (2003) study. This may have been due to task complexity, time-course factors, or some unknown combination of these and other issues.

References

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

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

Klatt, D. (1980) Software for a Cascade/Parallel Synthesizer. *Journal of the Acoustical Society of America* 67, 971-995.

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

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

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

McMurray, B. (in preparation). KlattWorks: a [somewhat] new systematic approach to formant-based speech synthesis for empirical research.

Ohala, J. J. and Feder, D. (1994). Listeners' normalization of vowel quality is influenced by 'restored' consonantal context. *Phonetica* 51, 111-118.

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

Repp, B. H. (1992). Perceptual restoration of a "missing" speech sound: Auditory induction or illusion? *Perception and Psychophysics* 51, 14-32.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General* 110, 474-494.

Samuel, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General* 125, 28-51.

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

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

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

Warren, R. M. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin* 96, 371-383.