Lexical Confusability and Degree of Coarticulation

REBECCA ANNE SCARBOROUGH
University of California, Los Angeles

1. Listener-motivated accommodations in speech

Speech is inherently communicative—"we speak to be heard in order to be understood" (Jakobson, Fant, & Halle, 1952). Thus, it stands to reason that speech production should be influenced by a speaker's desire to be understood. A speaker must produce a signal from which his listener can recover the intended message. If he is not sufficiently careful, communication will be unsuccessful. On the other hand, as long as a speaker remains intelligible to his listener, there is no reason that he cannot adjust his production to reduce the amount of effort he must expend as speaker. Lindblom (1990) has proposed a model of the interaction among the forces that shape speech, characterizing speech communication as a dynamic balance between speaker-oriented and listener-oriented forces, regulated by the communicative context. In other words, as factors in the communicative situation place extra demands on the listener, decreasing his chances of recovering the message, the speaker must adjust his pronunciation in order to produce clearer speech (referred to in Lindblom's model as "hyper-speech"). However, when conditions are favorable for communication, the speaker is free to conserve articulatory effort, producing reduced speech (or "hypo-speech").

Research has shown that speakers are sensitive to a number of different types of listener difficulties and make corresponding acoustic-phonetic accommodations. One early and easily confirmable observation of a listener-motivated speech accommodation was that people talk more loudly (Lombard 1911) and more slowly (Hanley & Steer 1949) in noisy environments than in quiet ones. Similarly, speech directed toward hearing-impaired listeners is slower and less phonologically reduced than normal conversational speech (Picheny, et al. 1986).

Importantly, these ostensibly listener-motivated accommodations (or at least the speech containing these accommodations) have also been shown to have an observable positive effect on intelligibility for listeners. Experiments have verified, for example, that speech produced in noise is more intelligible when presented at a constant speech-to-noise ratio than speech produced in quiet (e.g., Summers, et al. 1988). And Picheny, et al. (1985) present evidence of a substantial improvement in intelligibility for clear speech relative to normal conversational speech for hearing impaired listeners.

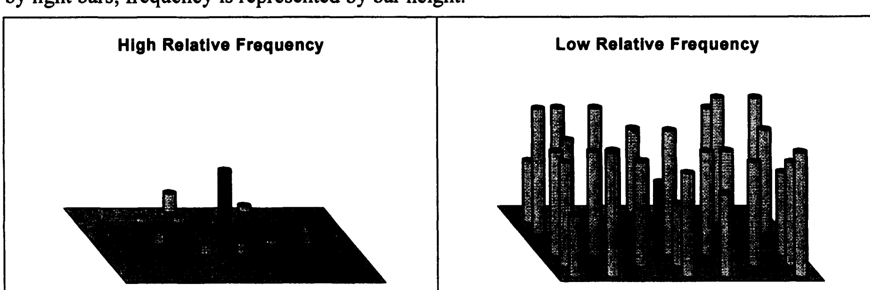
Factors internal to the structure of an interaction may also motivate speaker accommodations. Words that are less predictable from the conversational context are more intelligible when removed from their context and presented in isolation than are more predictable words (Lieberman 1963), and the first occurrence of a word in a narrative is more intelligible than the second occurrence of the same word when the words are heard out of their discourse context (Fowler & Housum 1987). These results have been interpreted as suggesting that new information and unpredictable information are spoken more clearly than old or predictable information because speakers know that listeners will be unable to rely on inferences from context to provide top-down cues to help them in lexical perception.

2. Lexical confusability

Factors at an even "lower", lexical level may also increase or decrease the likelihood of a word being correctly identified by a listener. For instance, more frequently occurring words are more quickly and accurately understood than less common ones (e.g., Gordon 1983; Glanzer & Eisenreich 1979). Intelligibility is further mediated by the number of words that are phonologically similar to a given word (i.e., the number of neighbors): the more neighbors a word has, the harder it is for a listener to identify (e.g., Luce 1986; Pisoni, et al. 1985). These two kinds of effects on the speaker-independent intelligibility of a word can both be understood as arising from the competition intrinsic in the process of lexical access.

Word identification involves discriminating among various entries in the mental lexicon. The difficulty of accessing a particular lexical item is related to the probability of correctly recognizing that item from among its lexical neighbors. Models of lexical access such as the Neighborhood Activation Model (Luce 1986; Luce & Pisoni 1998) propose that the frequency of a word and the number and frequencies of its competitor neighbors have various excitatory and inhibitory effects in lexical access that affect the activation of a target word relative to its competitors. The combined effects of frequency and neighborhood on the relative activation of a word can be represented by a ratio of the frequency of the word to the sum of the frequencies of all its neighbors (referred to here as relative frequency or *R*). The figure in (1) below illustrates high and low relative frequencies. In the high-*R* picture, the target word (which has a higher frequency than its few and infrequent neighbors) stands out, while in the low-*R* picture, the target word (with its relatively low frequency and many frequent neighbors) is obscured by its neighbors. With respect to intelligibility, then, high-*R* words are "easy" and low-*R* words are "hard". It is this sort of listener difficulty in word recognition that will be further explored in the current study.

(1) High vs. Low Relative Frequency: Target words are represented by dark bars, neighbors by light bars; frequency is represented by bar height.



That this notion of lexical confusability, measured by relative frequency, represents a real perceptual or intelligibility factor is verified by auditory lexical decision experiments showing faster response times to "easy" (high-*R*) words than to "hard" (low-*R*) words, indicating that "easy" words are more easily processed (Luce 1986; see also the current study). Such a measure also represents a real factor in production, as evidenced by the finding that "easy" words are produced with consistently greater vowel reduction than "hard" words (Wright 1997). In other words, the less easily processed words (the "hard" ones) are produced with more careful, acoustically distinct vowels, increasing the likelihood of intelligibility in precisely those cases where processing is more difficult (as evidenced by slower reaction times) and where a listener might otherwise be likely to misperceive.

3. Coarticulatory Modifications

The current study looks at another way speakers might accommodate the needs of listeners, investigating whether coarticulation is mediated by potential lexical confusability.

There are two competing hypotheses regarding the effect of lexical confusability on coarticulation. On the one hand, coarticulation introduces variability into the speech signal. Coarticulation is often viewed as the primary means by which speakers reduce their articulatory effort. But as coarticulation reduces speakers' effort, if it does, it may also reduce acoustic distinctiveness and hurt the ability of listeners to accurately or easily perceive the speech signal. From the point of view of the listener, then, coarticulation is just variability or noise that must be factored out of the signal in order to understand a speaker's message. Thus, in cases where speakers try to accommodate the difficulties of their listeners, it is predicted that they would increase their effort and decrease coarticulation, particularly in "hard" words.

On the other hand, though coarticulation may lead to a reduction of articulatory effort, it does not actually lead to a reduction of the amount of acoustic information contained in the speech signal. Coarticulation involves the overlapping of segments, spreading acoustic properties of one segment to another segment or segments. But these acoustic properties from different segments are not merged into a single, ambiguous acoustic stream. Rather, they are all present in the signal and available to listeners to perceive (Mattingly 1981). If the cues present in coarticulation can be used by listeners, then acoustic enhancement of certain segments by coarticulation could be beneficial to listeners in perceiving words. These considerations lead to the prediction that a speaker who is trying to be maximally intelligible would increase the degree of coarticulation in his speech, especially in "hard" words.

The current study investigates the effect of the listener-oriented constraint of lexical confusability on the production of coarticulation to determine whether coarticulatory modifications are used to increase speaker intelligibility in listener-directed speech. First, a production experiment seeks to resolve which of the two possible coarticulatory modifications hypothesized above, an increase or a decrease in coarticulation (if either), occurs in listener-directed speech. Then, a perception experiment tries to determine whether the coarticulatory modifications made in lexically confusable words is actually beneficial to listeners.

4. Production Experiment

Speakers were recorded producing words in two categories, high-*R* ("easy") and low-*R* ("hard"), which exemplify two types of coarticulation, carryover vowel-to-vowel coarticulation and anticipatory nasal coarticulation. The degree of coarticulation was then measured and compared across lexical confusability categories. The prediction was that coarticulation would be either reduced or exaggerated in "hard," or low-*R*, words relative to high-*R* words, reflecting speakers' knowledge about the effect of coarticulation on intelligibility.

4.1. Methods

4.1.1. Speech Materials

Tokens in the study included 40 monosyllabic words with codas containing nasals and 40 disyllabic words of the form CVC/i/. Within each set of tokens, half of the words were high-*R* and half were low-*R*. Lexical frequencies and the set of neighbors were determined based on the CELEX corpus (with adjustments for American English pronunciation) (Baayen, et al. 1995). For the purpose of the study, neighbors were considered to be all words differing from the target word by the addition, deletion, or substitution of a single phoneme (Greenberg & Jenkins 1964; Luce 1986). Relative frequency (*R)* was calculated by dividing the $\log_{10}$ frequency of the token word by the sum of the $\log_{10}$ frequencies of the word and all of its neighbors. Words were selected for inclusion in the study on the basis of having *R* values at the upper or lower end of the range of *R* values for words of the same type (i.e., CVNs or CVC/i/s). Additionally, mean log frequency and segmental context were balanced across the two relative frequency conditions for each type of coarticulation, and all words were of approximately equal, high subjective familiarity (Nusbaum, et al. 1984). For the purpose of comparison with the coarticulated vowel in the CVC/i/ words, a set of seven words exemplifying a canonical, non-coarticulated /i/ were also included in the list of tokens to be recorded. Four of these words were of the form C/i/C with no context for V-to-V coarticulation, and three were of the form C/i/C/i/ where the /i/s could only be coarticulated with themselves, yielding a form free from coarticulatory influence from other vowels. Two different randomized lists of the test words were created. The canonical /i/ words were included in a set of 12 practice words. (See Brown (2001) for a complete word list.)

4.1.2. Procedure

To encourage a listener-directed speech style, subjects participated in recording sessions in pairs, taking turns acting as speaker and listener. The speaker saw each word, one at a time, on a computer screen (not visible to the listener) and dictated a list to the listener, saying "The first word is X. The word after X is Y", etc. The speaker controlled the rate of presentation by pressing a button to advance to the next word. The procedure began with 12 practice tokens and continued until all 80 test tokens were uttered by the speaker. The roles of the two subjects were then reversed so that the subject who had been the speaker became the listener, and vice versa. The second speaker presented the words to the listener in the other random order. The procedure yielded two repetitions of each word for each speaker: one in utterance-final position, and the other in the "word after X" context. The data of seven native speakers of American English are analyzed here.

4.1.3. Measurements

The degree of nasal coarticulation in the CVN words was determined by measuring the nasality of the nasalized vowel from the acoustic signal. Nasalized vowels show the presence of an extra spectral peak at low frequencies, generally below the first formant, as well as a reduction in the amplitude of the first formant spectral peak. A measure of the relative amplitudes of the nasal peak and the first formant, A1-P0, where A1 is the amplitude of the first formant (as estimated by the amplitude of the peak harmonic closest to the expected F1) and P0 is the amplitude of the low frequency nasal peak (as estimated by the amplitude of either the first or second harmonic, depending on the speaker), has been found to be a reliable quantification of vowel nasality (Chen 1997). A1 and P0 were measured at the vowel midpoint and at the end of the vowel.

The degree of vowel-to-vowel coarticulation in the CVC/i/ words was determined by comparing the F1 and F2 of the second vowel (/i/) with the mean F1 and F2 values of the canonical, non-coarticulated /i/s for the same speaker. Since /i/ is higher and more front than any of the vowels which might be coarticulated with it, coarticulatory influence from any vowel would be realized as a raising of F1 and/or a lowering of F2. The first two formants of the canonical /i/ were determined by averaging the F1 and F2 of 18 /i/s in the practice words with no context for V-to-V coarticulation. Because V-to-V coarticulation is most visible in the region of the second vowel (V_2) closest to the first vowel (V_1), measurements for all vowels were made at V_2 onset and 20 ms into the vowel. All frequencies in Hertz were converted to the auditorily-scaled Bark scale, since the current study is concerned with the auditory salience of any acoustic differences that are found. Additionally, the transformation allows for the direct comparison of F1 and F2 differences. Differences between canonical /i/ and coarticulated /i/ were then calculated for both F1 and F2 at each measurement point.

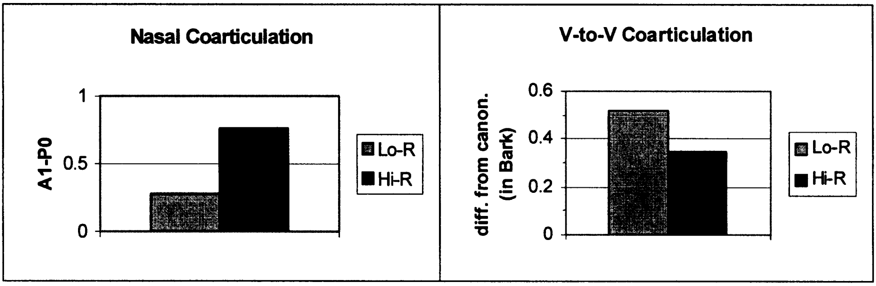
4.2. Results

Mixed repeated measures ANOVAs with nested within-subject factors of *R* (High vs. Low) and *Repetition* (First vs. Second) and a between-subjects factor of *RoleOrder* (A or B, indicating whether the subject was the first or second speaker in his session, with list randomization A assigned to the first speaker and B always assigned to the second) were performed separately on the A1-P0 values of the CVNs and the F1 and F2 differences of the CVC/i/s (averaged within each test category, for each measurement position).

For the nasal coarticulation items, a main effect of *R* was found at the vowel midpoint [$F(1,5)=11.77$, $p=.019$], reflecting that low-*R*, or "hard," words, have more nasalized vowels (i.e., lower A1-P0 values) than high-*R*, or "easy," words at the midpoint. This effect of *R* on nasality is summarized in the first graph in (2) below. No other main effects or interactions reached significance at any position, though low-*R* words also show smaller A1-P0 values than high-*R* words for five of the seven subjects at the end of the vowel. It was further observed that the A1-P0 values are smaller at the end of the vowel than at its midpoint, indicating that the vowels were more nasal closer to the nasal consonant.

For the V-to-V coarticulation items, there was again a significant main effect of *R* on F1 at V_2 onset [$F(1,5)=21.41$, $p=.006$], reflecting that the final vowels in low-*R* ("hard") words differ more from canonical /i/, presumably due to influence from V_1, than those in high-*R* ("easy") words. This effect of *R* is shown in the second graph in (2). No other main effects or interactions reached significance at V_2 onset or at 20 ms after onset, though F1 means at 20 ms show the same *R* pattern as that found at vowel onset. The F2 differences across high- and low-*R* words do not appear to vary systematically, however.

(2) Left: A1-P0 for Low-*R* vs. High-*R* CVN Words, at vowel midpoint. Right: F1 Difference between Canonical /i/ and Experimental /i/ for Low-*R* vs. High-*R* CVC/i/ Words, at vowel onset.



4.3. Discussion

The greater degree of nasality in low-*R* words relative to high-*R* words among CVNs indicates greater anticipatory nasal coarticulation in "hard", or more lexically confusable, words. Likewise, the greater deviation from canonical vowel values for F1 (at onset) indicates greater carryover coarticulation (at least on the height dimension) for "hard" CVC/i/ words. Thus, we see that "hard" words are produced with *more* coarticulation than "easy" words.

If we take the view that coarticulation is a sort of reduction (at very least, a strategy by which speakers reduce their overall effort) (e.g., Lindblom 1990), these results are quite surprising. The object of listener-directed speech is supposed to be to make speech more clear, i.e., less reduced, when there are greater demands on the listener. Wright's (1997) finding that "hard" words are produced with more dispersed vowels than "easy" words seems to underscore this principle of enhancing acoustic distinctiveness and minimizing reduction with respect to the demands of lexical confusability.

However, because the results of the current study indicate that speakers increase coarticulation in precisely those words that they otherwise try to make clearer, they suggest a very different view of coarticulation in which coarticulation is not in conflict with listeners' need to receive clear, distinct acoustic information. While coarticulation may in a sense diminish the acoustic distinctiveness of one segment (in that it makes the segment on which it is realized less canonical), it provides more extensive cues for another. It appears, then, that speakers produce low-*R* words with more coarticulation so that listeners can use the extra information contained in the coarticulation to help them perceive these "hard" words more easily.

If this story is correct, and coarticulation is increased in "hard" words in order to make these words maximally intelligible, gaps in the effect at certain measurement points or on certain dimensions are not expected. No effect was found, though, for nasality at the vowel endpoint in the case of nasal coarticulation or for F2 in the V-to-V case. Before we continue, then, we will briefly consider these cases. It is important to note that in these two instances, there is not a lack of coarticulation; in fact, there is as much coarticulation on the F2 dimension as there is for F1 and more nasal coarticulation at the endpoint than is found at other positions. What is lacking is simply an influence of *R*.

Looking first at A1-P0 at the endpoint in CVNs, it may be noted that, across all tokens, A1-P0 is significantly smaller at this point than at the midpoint [$t(28)=2.052$, $p<.0001$], indicating that the end of the vowel is more nasalized.

The lack of an effect of *R* can be straightforwardly interpreted as a ceiling effect. If the velum consistently reaches a fully low position before the closure for the nasal stop (toward the end of the vowel), it is not possible (or necessary) to differentiate the nasality of vowels in low-*R* words and high-*R* words by lowering the velum even further in the low-*R* words, since the vowels in both would already be maximally nasalized, produced with a maximally lowered velum.

With respect to F2 in the CVC/i/s, it may be noted that at both measurement positions, F2 differs as much from canonical in low-*R* ("hard") words as F1 does, suggesting the same amount of coarticulation for F1 and F2 in low-*R* words. However, there is also an F2 deviation from canonical in the high-*R* ("easy") words that is equal to that in the low-*R* words. It is possible, of course, that there is simply as much F2 coarticulation in "easy" words as in "hard" ones, even though we would not expect speakers to make any effort to adjust coarticulation in "easy" words. An F2 deviation, however, might also be attributable to reduction, an especially likely possibility for "easy" words, which undergo greater vowel reduction than "hard" words (Wright 1997). If F2 in "hard" words shows *R*-driven coarticulation, and F2 is lowered in the "easy" words due to vowel reduction, these two effects on F2 may balance each other out in the statistical analysis, hiding the effect of *R*. (Because there are more height (F1) contrasts than front-back (F2) contrasts to maintain in English, F1 is less likely to reduce than F2.) It would be helpful to examine the relative effects of reduction on F1 and F2 in "easy" words without a potential coarticulatory confound, though data on the contribution of the formants individually are not currently available.

5. Perception Experiment

We have supposed thus far that if speakers pronounce certain sounds differently depending on the relative frequency (*R*) of the word in which they occur, they must be doing so for the benefit of listeners, to make the words which are potentially the most difficult to perceive as easy as possible for the listener to understand. The purpose of the second experiment, then, is to investigate whether differences in pronunciation due to adjustments of the degree of coarticulation actually affect a listener's ability to perceive words. In the current experiment, listeners heard real words and nonsense words of two phonetic types: those which displayed normal coarticulatory effects and those in which coarticulatory effects were removed. Lexical decision reaction times, indicating a listener's facility in perceiving a word, from these two coarticulation conditions were compared. Based on the findings for low-*R* words in the production study, it was predicted that listeners would respond more quickly to tokens *with* coarticulation than to those *without*.

5.1. Methods

5.1.1. Stimuli

The set of stimuli included items that fall into three groups: test words, filler words, and non-words. The tokens of primary interest were the real test words, which were the same 40 CVNs and 40 CVC/i/s used in the production experiment. The set of filler words included 20 monosyllabic and 20 disyllabic forms, none of the form CVN or CVC/i/. The non-words included those corresponding to the test words, which were formed by changing the onset of the initial syllable, and those corresponding to the fillers, which were formed by changing any syllable onset or coda.

Each of the stimuli presented to subjects was produced by splicing together parts of two separate recordings. In some cases, the two recordings were different productions of the same item (referred to as "same-spliced"); in other cases, they were recordings of different items (referred to as "cross-spliced"). For each of the real test words, both a same-spliced version and a cross-spliced version were made. For the remaining items, only one version, either same-spliced or cross-spliced was made. The original components of the stimuli were produced by a female native speaker of American English. The speaker read a list containing the test words, fillers, non-words, and tokens resembling these items except that the segment inducing the relevant coarticulation was replaced, eliminating the context for coarticulation. For example, for the CVC/i/ item *chubby*, both *chubby* and *cheeby* were recorded. The tokens were spliced as shown in the figure in (3).

(3) Illustration of Splicing Scheme for V-to-V and Nasal Coarticulation

		Same-spliced	Cross-Spliced
V-to-V	Originals	**tʃʌb$_{ʌ}$i** **tʃʌb$_{ʌ}$i**	**tʃʌb$_{ʌ}$i** **tʃib$_{i}$i**
	Spliced stimulus	→ **tʃʌb$_{ʌ}$i** ←	→ **tʃʌb$_{i}$i** ←
Nasal	Originals	**bʌ̃n** **bʌ̃n**	**bʌ̃n** **bʌd**
	Spliced stimulus	→ **bʌ̃n** ←	→ **bʌn** ←

Because the reaction times to the pairs of real test words were to be compared (same-spliced versus cross-spliced), special care was taken to ensure that members of each pair were as similar as possible. The same sound file was used as the skeleton (the word without the critical vowel) for both versions, and the durations of the critical vowels were matched as closely as possible across the versions, differing by no more than 4 ms. (Further details on stimulus preparation are available in Brown (2001).)

All subjects heard the same 240 words and non-words; however, subjects were randomly divided into two groups with respect to which versions (same-spliced or cross-spliced) they heard of each word. Listeners in each group heard half of the real test words in each version, but the items that were heard in one version by one group were heard in their other version by the other group. Thus, each test token (i.e., a word in a particular splicing condition) was heard by half the subjects. The stimuli were heard in a random order.

5.1.2. Procedure

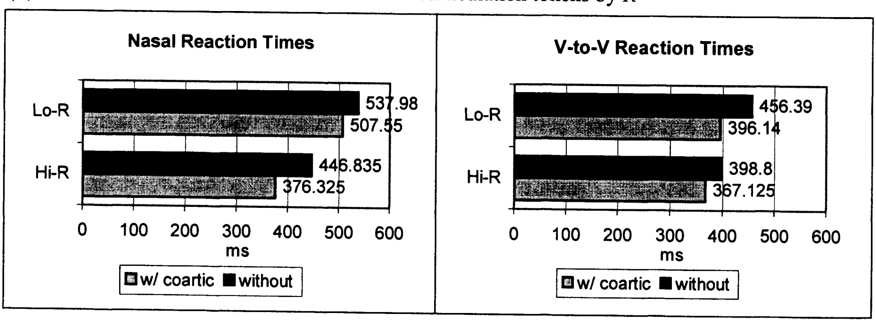
Subjects were instructed that they would hear a number of simple English words and some nonsense words and that their task was to decide quickly and accurately which were real words and which were non-words and to press a response button on a button box to indicate their choice. New trials were initiated 1 second after each button press. After 12 practice trials, responses ("word" or "non-word") and response latency times were recorded for 240 trials. Nineteen native speakers of American English, none of whom had participated in the first experiment, participated in this experiment.

5.2. Results

Overall, 91% of the 1520 trials involving the relevant real test words were correctly identified as words. Accuracy did not differ significantly between same- and cross-spliced tokens. In the analysis of reaction times, only data from real words that were correctly identified were considered. Furthermore, response times that did not fall within two standard deviations of the mean for each subject were also removed from the analysis.

Reaction times (averaged within each condition) for each subject were submitted to mixed repeated measures ANOVAs (one for each type of coarticulation, V-to-V and nasal) with nested within-subjects factors of Coarticulation (with vs. without, represented by the same- vs. cross-spliced tokens) and *R* (high vs. low) and a between-subjects factor of List (A or B, indicating which set of words subjects heard in which condition). For both types of coarticulation, there is a significant main effect of *R* [Nasal: $F(1,17)=61.57$, $p<.0001$], [V-to-V: $F(1,17)=16.7$, $p=.0008$], reflecting that low-*R* words were, as expected, harder (i.e., they took longer) for subjects to verify. The hypothesized main effect of Coarticulation was also found to be significant for both coarticulation types [Nasal: $F(1,17)=28.54$, $p<.0001$], [V-to-V: $F(1,17)=5.77$, $p=.028$], reflecting the fact that words *with* coarticulation (i.e., same-spliced words) were more quickly verified than words with no coarticulation (i.e., cross-spliced words). The reaction time data are summarized in (4).

(4) Reaction Times to with- vs. without-coarticulation tokens by *R*



5.3. Discussion

The results from the lexical decision experiment showing a significant effect of Coarticulation support the hypothesis that the presence of coarticulation facilitates lexical perception. Interestingly, these results obtain both when coarticulation provides cues for an upcoming segment (as in the nasal case) *and* when it reinforces cues following a segment (as in the carryover V-to-V case).

Earlier studies have shown that listeners are able to attend to coarticulatory information when they are asked to make explicit, fine, phonetic discriminations (e.g., Beddor, et al. 2001), and that they are sensitive to the inappropriateness of certain bits of coarticulatory information (e.g., Fowler 1984). The current study extends these findings, allowing for broader generalizations regarding listeners' abilities to use information contained in coarticulation. First, it examined a relatively more linguistically-natural task. Listeners' reactions to appropriately coarticulated strings were compared to coarticulatorily-neutral strings, rather than to instances of inappropriate coarticulation. Because inappropriate coarticulation (e.g., vowel nasality before an oral consonant) is clearly unnatural, it might be expected to be harder to respond to on the basis of a general listener dispreference for unnaturalness, whether or not the presence of appropriate coarticulatory cues in the other condition contributed to improved reaction times. In the current study, however, the inappropriate coarticulation condition was replaced with a neutral condition, in which coarticulation was simply removed. Because coarticulation naturally varies in degree, tokens that lack coarticulation should not sound completely unnatural; rather they should sound like tokens produced with the smallest possible degree of coarticulation. Furthermore, the current study looked at listeners' reactions to coarticulation as they were processing real words (as opposed to nonsense phonetic strings). The current findings, then, suggest that listeners are not only able to extract phonetic information from coarticulation, but that they actually do make use of coarticulatory information (at least under the circumstance of timed lexical decision) when they perceive real words.

6. Summary and General Discussion

The purpose of the current study was to investigate the effect of potential lexical confusability on the production of coarticulation. We asked whether speakers would try to accommodate listeners' difficulties in spoken word recognition caused by lexical confusability by adjusting the amount of coarticulation in particularly confusable words. And we looked at the effect of the presence of coarticulation on lexical identification to see whether the accommodations made by speakers would actually be beneficial to listeners. It was found that more lexically confusable words, those with low frequencies relative to their lexical neighbors, were produced with a greater degree of coarticulation than less confusable words for both types of coarticulation examined (carryover vowel-to-vowel and anticipatory nasal coarticulation). It was also shown that listeners were able to more quickly identify words containing coarticulation than words in which the coarticulation had been eliminated.

Taken together, the results from these two experiments provide a coherent story within the framework of listener-directed speech accommodations. Listeners are able to recognize words with more coarticulation more quickly and easily than those without, so speakers adjust their pronunciation of harder, more confusable words to exhibit a greater degree of coarticulation in order to facilitate lexical recognition. To explain this phenomenon, coarticulation can be thought of as a process that spreads the properties of a segment, causing them to overlap in time with cues from neighboring segments. (For a discussion of this sort of spreading process in phonology, see Kaun (1995) in which certain instances of phonological assimilation are characterized as spreading to enhance features.) Listeners are able to compensate for coarticulation, correctly attributing the acoustic consequences of coarticulation to their source (e.g., Beddor & Krakow 1999; Mann & Repp 1980). And they are also able to use coarticulatory information to identify or predict other portions of the signal (e.g., Fowler 1984; Beddor & Krakow 1999). Thus, increased coarticulation in "hard" words (at least to a certain degree) can provide extra cues for the coarticulatory source segment without interfering with the cues for the segment on which they occur. This is the effect that emerges from the present study.

The current study, then, suggests that speakers have some knowledge of lexical confusability or relative frequency that leads them to identify certain words as potentially difficult (with respect to lexical access) and that triggers them to make various communicatively-motivated accommodations. It demonstrates that one such accommodation involves an increase in the degree of coarticulation in "hard" words. And it suggests that coarticulation is more than just a type of reduction that allows speakers greater ease of articulation. Some coarticulation (the amount that we see in high-*R* words) may in fact facilitate articulation (and may even be necessary for articulation). But beyond this, coarticulation (for example, in the amount that we see in low-*R* words) serves an explicitly listener-oriented, communicative purpose, providing additional cues that facilitate lexical perception.

References

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995) The CELEX Lexical Database (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [distributor].

Beddor, P. S. & Krakow, R. A. (1999) Perception of coarticulatory nasalization by speakers of English and Thai: evidence for partial compensation. *JASA*, 106, pp. 2868-2887.

Beddor, P. S., Krakow, R. A., & Lindemann, S. (2001) In E. V. Hume & K. Johnson (Eds.), *The Role of Speech Perception Phenomena in Phonology.* Academic Press.

Brown, R. A. (2001) *Effects of Lexical Confusability on the Production of Coarticulation.* UCLA master's thesis.

Chen, M. Y. (1997) Acoustic correlates of English and French nasalized vowels. *JASA*, 102, pp. 2360-2370.

Fowler, C. A. (1984) Segmentation of coarticulated speech in perception. *Perception & Psychophysics*, 36, 359-368.

Fowler, C. A. & Housum, J. (1987) Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction. *JML*, 26, 489-504.

Glanzer, M. & Eisenreich, S. L. (1979) Structure and search of the internal lexicon. *J. Verbal Learning and Verbal Behavior*, 18, 381-398.

Gordon, B. (1983) Lexical access and lexical decision: mechanisms of frequency sensitivity. *J. Verbal Learning and Verbal Behavior*, 18, 24-44.

Greenberg, J. H. & Jenkins, J. J. (1964) Studies in the psychological correlates of the sound system of American English. *Word*, 20, 157-177.

Hanley, T. D. & Steer, M. D. (1949) Effect of level of distracting noise upon speaking rate, duration and intensity. *J. Speech and Hearing Disorders*, 14, 363-368.

Jakobson, R., Fant, C. G. M., & Halle, M. (1952) Preliminaries to speech analysis: the distinctive features and their correlates. *Technical Report, no. 13.* Cambridge, MA: Acoustics Laboratory, MIT.

Kaun, A. R. (1995) *The Typology of Rounding Harmony: An Optimality Theoretic Approach.* UCLA dissertation.

Lieberman, P. (1963) Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172-187.

Lindblom, B. (1990) Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modeling*, pp. 403-439. Netherlands: Kluwer Academic Publishers.

Lombard, E. (1911) Le signe de l'élévation de la voix. *Annales des maladies de l'oreille, du larynx, du nez et du pharynx*, 37, 101-119.

Luce, P. A. (1986) Neighborhoods of words in the mental lexicon. *Research on Speech Perception Progress Report, No. 6.* Speech Research Lab., Psych. Dept., Indiana University.

Luce, P. A. & Pisoni, D. B. (1998) Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*, 19, 1-36.

Mann, V. & Repp, B. (1980) Influence of vocalic context on perception of the [š]-[s] distinction. *Perception & Psychophysics*, 28, 213-228.

Mattingly, I. G. (1981) Phonetic representation and speech synthesis by rule. In T. Meyers, J. Laver, & J. Anderson (Eds.), *The Cognitive Representation of Speech,* pp. 415-420. Amsterdam: North-Holland Publishing Company.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984) Sizing up the Hoosier mental lexicon: measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report, No. 10.* Speech Research Lab., Psych. Dept., Indiana University.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985) Speaking clearly for the hard of hearing I: intelligibility differences between clear and conversational speech. *J. Speech and Hearing Research*, 28, 96-103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986) Speaking clearly for the hard of hearing II: acoustic characteristics of clear and conversational speech. *J. Speech and Hearing Research*, 29, 434-446.

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985) Speech perception, word recognition, and the structure of the lexicon. *Speech Communication*, 4, 75-95.

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988) Effects of noise on speech production: acoustic and perceptual analyses. *JASA*, 84, 917-928.

Wright, R. (1997) Lexical competition and reduction in speech: a preliminary report. *Research on Spoken Language Processing Progress Report, No. 21.* Speech Research Lab., Psych. Dept., Indiana University.

Rebecca Scarborough
3125 Campbell Hall
UCLA Linguistics Department
Los Angeles, CA 90095

rab@humnet.ucla.edu