

Effects of Anticipatory Coarticulation on Lexical Access

Stephen J. Tobin (stephen.tobin@uconn.edu)

Department of Psychology, University of Connecticut, Storrs, CT 06269-1020 USA
Haskins Laboratories, 300 George St., New Haven, CT 06510 USA

Pyeong Whan Cho (pyeong.cho@uconn.edu)

Department of Psychology, University of Connecticut, Storrs, CT 06269-1020 USA
Haskins Laboratories, 300 George St., New Haven, CT 06510 USA

Patrick M. Jennett (pjennett@engr.uconn.edu)

Cognitive Science Program, University of Connecticut, Storrs, CT 06269-2054 USA

James S. Magnuson (james.magnuson@uconn.edu)

Department of Psychology, University of Connecticut, Storrs, CT 06269-1020 USA
Haskins Laboratories, 300 George St., New Haven, CT 06510 USA

Abstract

One of the most challenging unsolved problems in cognitive science is lack of invariance in spoken language. We take the view that variability due to coarticulation is systematic and beneficial. Several recent eye tracking experiments have demonstrated listeners' sensitivity to local coarticulatory cues between adjacent phonemes. We examined sensitivity to longer-range, anticipatory vowel-to-vowel coarticulation, which can spread across multiple syllables. Using a variant of the Visual World eye tracking paradigm (Tanenhaus et al., 1995), we conducted the first on-line test of whether lexical access is sensitive to such subtle, long-range cues, and whether the impact of such cues is modulated by the coarticulation resistance of intervening segments. Lexical access was delayed when misleading anticipatory coarticulation was available in cross-spliced materials. This significantly extends the nature and temporal range of subcategorical cues known to influence on-line sentence comprehension, and demonstrates that lexical access is simultaneously constrained by information at multiple temporal grains.

Keywords: Coarticulation; anticipation; garden path; eye tracking.

Introduction

One of the hardest unsolved problems in cognitive science is *lack of invariance* in speech. There is a many-to-many mapping between acoustics and percepts, such that the same acoustic information can map to different speech sounds, while different acoustic information can map to the same speech sounds (depending on phonetic context, speaking rate, physical or indexical characteristics of talkers, etc.). This is true of production and perception even for clearly articulated segments and syllables (Ladefoged & Broadbent, 1957; Liberman, Delattre & Cooper, 1952; Peterson & Barney, 1952). The problem is compounded in mapping to words and beyond in conversational speech, where even more variation occurs. For example, Hawkins (2003) describes radical changes in the acoustics of the message "I do not know" in a progression from careful speech to casual speech ("I dunno", and even more reduced forms). The puzzle, then, is how we reliably map acoustics to words despite (or perhaps with the aid of) all this variation.

In order to make progress in studying spoken word recognition, psycholinguists have made the temporary simplifying assumption that the input to word recognition can be approximated by a phonemic transcription (as though this were the product of a speech perception mechanism). This allows one to sidestep the lack of invariance problem and related complications due to *coarticulation*. Coarticulation refers to the fact that the articulatory gestures of adjacent and even nonadjacent segments overlap, and therefore, so do their acoustic realizations. That is, as you produce one speech sound, you are simultaneously preparing your articulators for upcoming segments, and still experiencing effects of preceding articulations. Coarticulation is often viewed as destructive, as in Hockett's (1955) metaphor of a wringer squishing together a line of easter eggs on a conveyor belt, and a major contributor to lack of invariance.

Even when a scientist is cognizant of the fact that the phonemic input assumption is almost certainly incorrect, and she explicitly considers it provisional (until we solve the lack of invariance problem at the phonological level), it has the potential to hide constraints on word recognition (Magnuson, 2008). For example, Salverda, Dahan and McQueen (2003) reported that listeners use subtle prosodic cues (e.g., vowel duration) to anticipate word length and constrain lexical competition. They tracked eye movements as subjects followed spoken instructions to click on pictures on a computer display. When initial vowel duration was consistent with a bisyllabic word, subjects immediately began looking preferentially at items with bisyllabic names.

A phonemic transcription also abstracts away from coarticulatory information, which can specify qualities of upcoming segments. Dahan, Magnuson, Tanenhaus, and Hogan (2001) demonstrated that listeners are extremely sensitive to such information. They cross-spliced words (e.g., *neck* and *net*) to provide misleading coarticulatory cues to final consonants. Using the visual world paradigm, they found fast, robust effects of such mismatches on lexical activation and competition.

These examples are inconsistent with suggestions that coarticulation has a destructive impact on phonetic

information. Instead, coarticulation systematically provides anticipatory and redundant cues that afford rapid information transmission in speech. This optimistic view, that coarticulation is lawful and informative (Elman & McClelland, 1986; Fowler, 1986), is consistent with evidence that listeners compensate for coarticulation, taking into account the predictable structure of an ongoing speech event at the gestural (Fowler, 1980; Browman & Goldstein, 1992), phonological (Gow, 2001), lexical (Ganong, 1980; Magnuson et al., 2003), and sentential (Gaskell & Marslen-Wilson, 2001) levels.

All of these examples involve cases of local coarticulation, that is, coarticulation between adjacent segments. Subsequent work on speech production has revealed the existence of long-range coarticulation, in some cases spanning multiple segments or even syllables (Heid & Hawkins, 2000; Recasens, 1984; West, 2000). This raises the possibility that listeners have even richer information at their disposal -- cues specifying qualities of upcoming sounds even several segments in advance. However, these effects are subtle and subject to strong constraints. Among these constraints is *coarticulation resistance*, a finding of Bladon and Al-Bamerni (1976), who observed that intervening consonants could modulate the effects of vowel-to-vowel coarticulation. Specifically, light, palatal [l]s reduced coarticulation between surrounding vowels in comparison to dark, velarized [ɫ]s. The articulatory and perceptual effects of coarticulation resistance has been investigated in some detail (Fowler, 2005; Fowler & Brancazio, 2000; Recasens & Espinosa, 2009). Consonants with high coarticulation resistance prevent coarticulation between surrounding vowels. Typically, high coarticulation resistance consonants involve tongue body or tongue tip articulations (e.g., [t], [d]) and/or fine motor control (e.g., [s], [z]). That is, strong constraints on tongue tip or tongue body damp long-range, vocalic coarticulation. Low coarticulation resistance consonants allow coarticulatory information to spread further, as they do not involve the tongue body or tip and do not require particularly fine motor control (e.g., [p], [b], [f], [v]).

The current study is the first on-line study, to the best of our knowledge, to examine whether lexical access is sensitive to long-range coarticulatory information, and whether the impact of such information is modulated by the coarticulation resistance of intervening segments. If so, this will represent a substantial increase in the amount of detail listeners are known to use in order to constrain speech perception and word recognition.

Experiment

We examined whether coarticulatory effects would influence lexical access by manipulating two factors. The first was (*Coarticulatory*) *Match*. At the Match level, two instances of one utterance (e.g., "pick up a pole") were cross-spliced after the word "a". In the Mismatch condition, two utterances with different final vowels were

cross-spliced (e.g., "pick up a pole" and "pick up a pail"). This provides a potentially more powerful window on sensitivity to long-range coarticulation; the Mismatch stimuli should slow lexical access of the final target word, since the coarticulation is consistent with another word, which should compete more strongly for lexical access.

The second factor was (*Coarticulation*) *Resistance*. After low coarticulation resistance consonants such as [p], the full vowel in the final word in the utterance, "pick up a pole", is likely to influence the realization of the reduced vowels in "up" and "a." In contrast, [t] is high coarticulation resistant, so anticipatory coarticulation from the vowel would be less likely if the final word were "toll." Therefore, we selected words beginning with High Resistance ([t, s, S]) or Low Resistance ([p, f]) segments.

Predictions

We used a variant of the visual world paradigm with two printed words as response choices (e.g., *pole*, *pail*), as subjects heard sentences like, "pick up a pole". Our first question is exploratory: whether and when subjects might begin to favor one word based on anticipatory coarticulatory cues. Given a low resistance consonant, it is possible that subjects could begin to pick up information about the final vowel as early as the vowel in "up." When Match and Resistance are crossed, an interaction is predicted: the effect of Match should be most apparent at the Low level of Resistance.

Methods

Participants Thirty-one undergraduate students at the University of Connecticut participated in this experiment for course credit. All were native English speakers with normal or corrected-to-normal vision and reported normal hearing.

Table 1: Low and High Coarticulation resistance items. Numbers indicate quadruple set membership.

Low Coarticulation Resistance Pairs		High Coarticulation Resistance Pairs	
1.pail,pole	6.fake,folk	1.tail,toll	6.sake,soak
2.pea,porch	7.fail,foal	2.tea,torch	7.sail,sole
3.paste,post	8.feel,fault	3.taste,toast	8.seal,salt
4.pan,pool	9.feet,fog	4.tan,tool	9.seat,sauce
5.peak,pork	10.feed,ford	5.teak,torque	10.seed,sword
	11.field,fall		11.shield,shawl
	12.feet,fort		12.sheet,short

Materials Twelve quadruples of words were chosen for the study. Each was composed of a pair of words starting with a low coarticulation resistance consonant (e.g., *pail/pole*), and a pair starting with a high coarticulation resistance consonant (e.g., *tail/toll*; see Table 1 for the full set). A number of constraints were observed in the selection of these quadruples. (1) The phonemes /p/ or /f/ were used as low coarticulation resistance consonants, and /t/, /s/, or /ʃ/ were used as the high coarticulation resistance consonants. (2) Highly discriminable front or back vowels were used to maximize acoustic differences

between words and also allow for maximal acoustic difference in anticipatory vocalic coarticulation. The front vowels used were: /i/, /eɪ/ and /æ/. The back vowels were: /əʊ/, /ɑ/, /o/ and /u/. (3) When possible, we used the same final consonant in all words in a quadruple, while also varying length and frequency as little as possible. These constraints had to be relaxed in a few cases in order to find enough items. However, the most critical portion of any noun in our design is the initial consonant and vowel, which constrain the potential for long-range coarticulation. ANOVAs confirmed that items did not differ reliably in any of these characteristics.

Each word in a quadruple was recorded with the same sentence frame (e.g., “Pick up a”). This particular sentence frame was selected to be as naturalistic as possible, while also containing neutral vowels (/ə/ in “up” and “a”), thereby maximizing the chance of observing long-range coarticulatory effects. A male, native English speaker recorded all of the sentences at a moderately fast speaking rate, which produced observable long-range coarticulation. The auditory stimuli were recorded and presented in 16-bit resolution at a 44.1 kHz sampling rate.

The spoken sentences were all cross-spliced at the onset of the noun-initial consonant. In the Match condition, two tokens of the same recording were spliced together, to ensure that any effects in the Mismatch condition were not due to artifacts of cross-splicing. In the Mismatch condition, the carrier phrase from one recording (e.g., the “pick up a” portion of “pick up a pail”) was cross-spliced with the noun from another recording (e.g., the “pole” of “pick up a pole”). Average durations were 121 ms for “pick”, 171 ms for “up a”, and 317 ms for nouns.

Table 2: Mean F2-F1 (Hz) for *up* and *a* by condition.

	Coarticulation Resistance			
	Low		High	
	<i>up</i>	<i>a</i>	<i>up</i>	<i>a</i>
Front Vowels	789	793	811	1064
Back Vowels	772	612	808	938
ΔV	17	181	3	126

Acoustic analysis The formants of vowels of ‘up’ and ‘a’ of each target sentence were measured. Vowel formant center frequencies were measured using LPC and FFT spectra with reference to a wideband spectrogram. Measurements were made at the most stable portion of the middle of the vowel. Following Ladefoged (1993) we used F2-F1 (second formant - first formant) as a measure of vowel backness. The results are summarized in Table 2, collapsing over Front (/i, eɪ, æ/) and Back (/əʊ, ɑ, o, u/) vowels. F2 and F1 are more widely spaced for front than back vowels, so F2-F1 should be greater for front vowels. High Resistant consonants should yield smaller vowel backness differences than Low Resistant consonants. All of these differences were observed in the mean F2-F1 values at both ‘up’ and ‘a’. Thus, we were successful in

providing long-range anticipatory coarticulation cues in the materials.

Procedure Participants were seated at a comfortable distance from a computer screen (approximately 60 cm). Eye movements were monitored with an SR Research EyeLink 1000 desktop-mounted (remote) system, sampling at 500 Hz. Spoken sentences were presented to participants through headphones.

Each trial started with a drift correction procedure (participants briefly fixated a central dot). Then, a central fixation cross appeared. Participants clicked on it to begin the trial. When the cross was clicked, the members of a word pair appeared on the screen, one on the left and one on the right, with target and distractor position counter-balanced and pseudo-randomized. (We did not use pictures because we were unable to find enough highly imageable words meeting our phonological constraints; see McMurray, Tanenhaus & Aslin, 2009, and Huettig & McQueen, 2007, for precedents of using printed words in the visual world paradigm to obtain fine-grained time course measures of speech perception and spoken word recognition.) After a delay of 500 ms, the spoken sentence was presented over headphones. Participants were instructed to click on the final word of each sentence. The trial ended when the participant clicked on a printed word.

Each participant was presented with all 48 experimental trials and 44 filler trials. Twenty-two fillers consisted of rhyming pairs (to direct attention away from the onset similarity of critical items), and the remaining twenty-two were non-rhyming pairs. Half of each of these sets were presented with an auditory sentence cross-spliced to produce mismatching coarticulation, and the other half were spliced with another token of the same sentence, using exactly the same procedure as for the experimental stimuli (due to space constraints, we do not report analyses of filler items here). Experimental and filler trials were presented in random order following four (filler) practice trials. Half the experimental trials were presented in the Match condition and half were presented in the Mismatch condition. Half the trials in the Match and Mismatch conditions were Low Coarticulation Resistance words and the others were High. Thus, participants were given 12 trials in each of the four possible conditions. The word pairs (*pail/pole*) were counterbalanced between participants, such that a participant heard one of the pair (*pail*) in the Match condition, and the other (*pole*) in the Mismatch condition. Across participants, each word pair was presented the same number of times in Match and Mismatch conditions, and each word appeared equally often on the left or right.

Results

Data from two participants was excluded from analyses because of poor eye tracker calibration. Mean accuracy was at least 0.99 in all conditions. Figure 1 shows the average time course of target and competitor fixation proportions (at 20 ms intervals). Qualitatively, one

Table 3: Growth curve analysis of competitor fixation proportions. See text for details.

	Model fit			Match			Coarticulation Resistance			Match x CR		
	-2LL	ΔD	p	Estimate	t	p	Estimate	t	p	Estimate	t	p
Base	5996.8	-	-	-	-	-	-	-	-	-	-	-
Intercept	6013.8	17.1	0.001	0.074	4.39	<0.001	0.011	0.63	0.528	-0.058	-2.43	0.017
Linear	6018.9	5.1	0.166	-0.138	-2.16	0.031	-0.043	-0.68	0.500	0.130	1.44	0.149
Quadratic	6097.1	78.2	<0.001	-0.122	-7.57	<0.001	-0.001	-0.04	0.971	0.067	2.94	0.003
Cubic	6151	53.9	<0.001	0.111	6.92	<0.001	0.033	2.04	0.041	-0.082	-3.60	<0.001

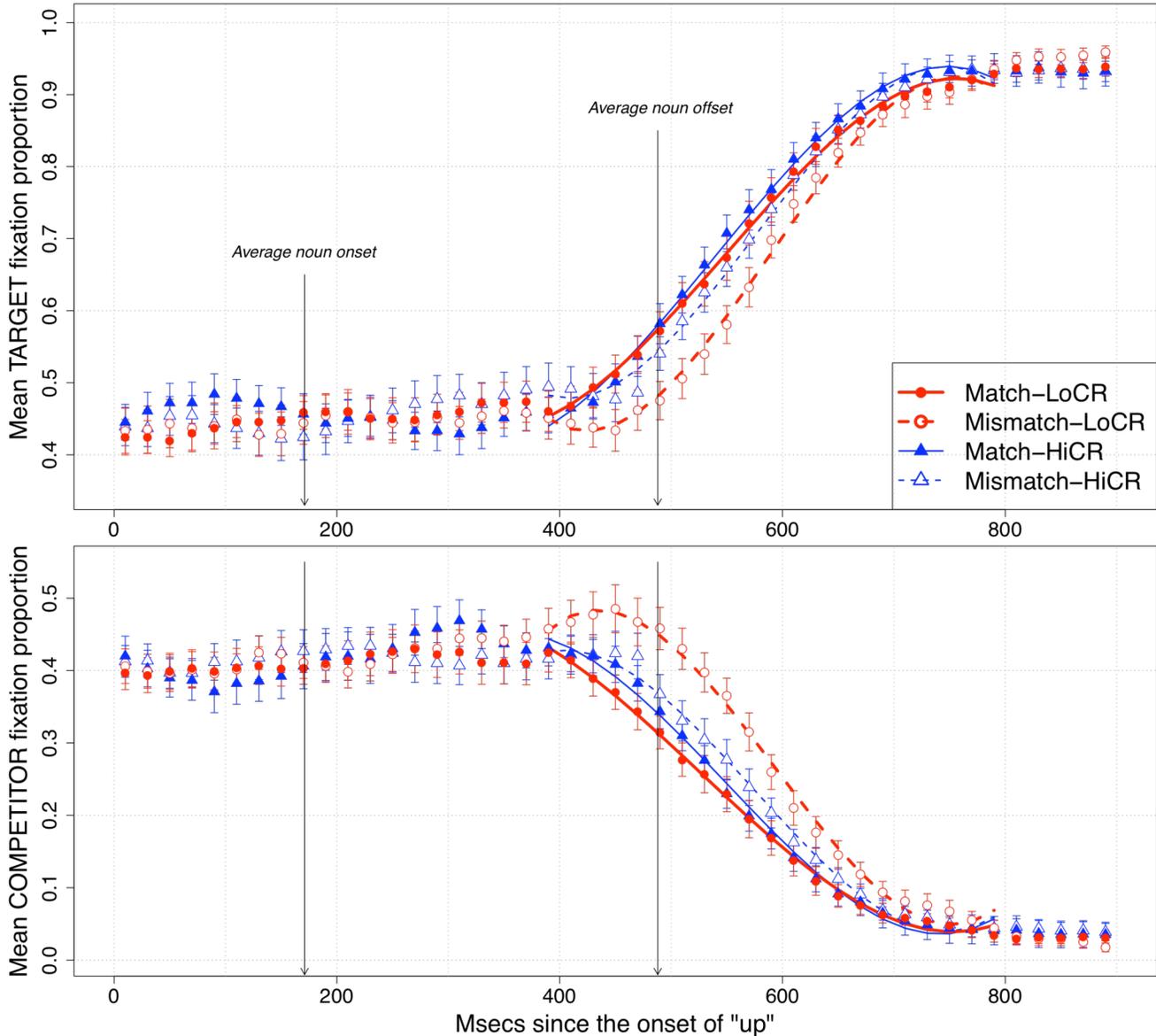


Figure 1: Mean fixation proportions to targets (top) and competitors (bottom). Symbols show observed data. Lines show growth curve fits for the 371-800 ms analysis window. Note different y-axis ranges.

condition stands out. The target is fixated most slowly in the Mismatch, Low Resistance condition, and a complementary increase in competitor fixations is also observed. The effect is not apparent until midway through the noun; however, the difference can only be due to coarticulatory detail available prior to word onset, since

the signal in that condition was identical to that in the Match, Low Resistance condition from noun onset onward. There was also a slight trend towards an effect of Match at the High level of Coarticulation Resistance.

For the statistical analyses, we applied *growth curve analysis* (GCA; Mirman, Dixon, & Magnuson, 2008).

GCA is a variant of multi-level modeling that fits orthogonal power polynomial terms to over-time data. Conceptually, curves are fit to subject mean proportions at the lowest level of condition combinations (e.g., the Match level at the Low level of Resistance), and GCA assesses whether the curve parameters differ. Notably, this approach is dynamically consistent: the average of subject-level fits is equivalent to the fit to averaged data.

It is clear from Figure 1 that there were no differences between conditions until midway through the noun. Therefore, we constrained the GCA to a window beginning 200 ms after noun onset, the approximately earliest point where noun-driven changes in fixation proportions could be observed (371 ms after "up" onset), and ending at target asymptote (800 ms).

While any order polynomial can be used, the first three terms are conceptually easiest to link to visual world data. The *intercept* is recentered, such that it is analogous to mean proportion in the analysis window, and so directly analogous to an ANOVA on mean fixation proportion (indeed, although standard ANOVAs are less powerful than the multi-level modeling afforded by GCA, ANOVAs on mean proportion in this analysis window converge with the GCA intercept analyses). The *linear* term is the mean slope over the analysis window, the *quadratic* term reflects bowing of the primary curve inflection, and the *cubic* term capture inflections at the tails (necessary for fitting the s-like curves here).

Table 4: Mean competitor fixation proportions in the 371-800 ms analysis window by condition.

	<i>Coarticulation Resistance</i>	
	Low	High
Match	0.198	0.209
Mismatch	0.271	0.225

Because effects on targets and competitors are logically complementary (since these categories represent the two primary objects of fixation), we will present just the competitor analysis. The analyses are summarized in Table 3, with main effects of Match on intercept (greater mean proportion in Mismatch than Match conditions), linear slope and quadratic curvature (due to more negative slope and greater curvature in Mismatch conditions that follows from the greater lag preceding the drop-off in competitor fixations), and the cubic term (due to the early initial rise in the Mismatch, Low condition). There were no main effects of Coarticulation Resistance, but there were interactions of Match and Coarticulation Resistance on intercept, quadratic, and cubic terms. The intercept interaction is crucial (and space does not permit discussion of the other interactions); simple effects analyses reveal a significant effect of Match at Low Resistance ($t=19.5$, $p<0.001$) but not at High ($t=1.5$). Relevant means are presented in Table 4. Thus, the predicted interaction was observed: the effect of

Mismatch was only reliable in the context of Low Coarticulation Resistance consonants.

Discussion

This study addresses the question of whether long-range coarticulatory information influences the time course of lexical activation and competition. We systematically varied (a) the Coarticulation Resistance (high vs. low) of the onset consonant of a monosyllabic target noun and (b) Coarticulatory "Match", i.e., whether the long-range anticipatory coarticulation matched or mismatched the vowel of the target noun that was ultimately heard.

While we observed trends associated with Coarticulatory Match and Resistance, the effects were driven largely by a difference in one condition: responses were slowed in the Mismatch, Low Resistance condition (though there was also a slight trend toward an effect of Mismatch at High Coarticulation Resistance).

The fact that the influence of anticipatory coarticulation was most apparent at low Coarticulation Resistance is consistent with phonetic analyses of long-range coarticulation, as those effects are simply more likely to propagate over segments (like [p]) that do not impose strong constraints on the position of the tongue tip or tongue body. The fact that influences of anticipatory coarticulation were most apparent in the Mismatch conditions is consistent with our expectation that these would be subtle effects and that misleading cues might be required to elicit detectable changes in lexical access (cf. Dahan, Magnuson, Tanenhaus & Hogan, 2001).

Notably, the effects do not emerge until midway through the final noun. This is surprising, as these effects must be due to anticipatory coarticulatory effects, since the nouns in the Match and Mismatch conditions were identical; those conditions differed only the "pick up a" portion of the instruction. It may be that anticipatory effects are detected as they occur, but require combination with confirmatory bottom-up evidence before they have a detectable impact. If front or back vowel qualities are detected on the vowel in 'up', this indicates that such a vowel is forthcoming, but it may still be several syllables away. This could prime appropriate phonological representations without driving strong lexical activation.

We do not mean to imply that lexical access initially proceeds based on local, bottom-up cues and other constraints are integrated after a delay (cf. Swinney, 1979). Instead, we note that when constraints are *gated* by bottom-up information but are integrated continuously, weak effects can appear to be late effects (Dahan, Magnuson & Tanenhaus, 2001; Shillcock & Bard, 1993).

Even though the effects we observed were relatively late and relatively modest, we did find a reliable influence of long-range coarticulatory anticipation. This reveals how extraordinarily sensitive listeners are to the rich sea of subphonemic details (some strong, some subtle) in which the islands of stability we describe as "phonemes" are embedded. This reinforces the view that coarticulation and other sources of variability in speech are not noise or

problems listeners must overcome. Rather, variability is largely lawful, enabling rapid rates of information transmission (Elman & McClelland, 1986; Fowler, 1986) due to local and anticipatory constraints it provides at multiple temporal scales.

Acknowledgments

Supported in part by NIH HD-40353 to Haskins Labs, and NSF grants CAREER 0748684 and 0642300 to JSM.

References

- Bladon, R. A., & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *J. of Phonetics*, 4, 137-150.
- Browman, C., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155-180.
- Dahan, D., Magnuson, J. S., and Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, 42, 317-367.
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: evidence for lexical competition. *Language & Cog. Processes*, 16, 507-534.
- Elman, J. L. & McClelland, J. L. (1986). Exploiting the lawful variability in the speech wave. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability of speech processes*. Hillsdale, NJ: Erlbaum.
- Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8, 113-133.
- Fowler, C. A., (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3-28.
- Fowler, C. A. (2005). Parsing coarticulated speech in perception: Effects of coarticulation resistance. *Journal of Phonetics*, 33(2), 199-213.
- Fowler, C. A., & Brancazio, L. (2000). Coarticulation Resistance of American English Consonants and its Effects on Transconsonantal Vowel-to-Vowel Coarticulation. *Language and Speech*, 43 (1), 1-41.
- Ganong, W.F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception & Performance*, 6, 110-125.
- Gaskell, M. G. & Marslen-Wilson, W. D. (2001). Lexical ambiguity and spoken word recognition: bridging the gap. *Journal of Memory and Language*, 44, 325-349.
- Gow, J. D. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, 45, 133-159.
- Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, 31, 373-405.
- Heid, S., & Hawkins, S. (2000) An acoustical study of long-domain /t/ and /l/ coarticulation. *Proc. 5th Seminar on Speech Production: Models and Data*. (ISCA). Kloster Seeon, Bavaria, Germany. 77-80.
- Hockett, C. F. (1955). A Manual of Phonology. *International Journal of American Linguistics (Memoir II)*. Baltimore: Waverley Press.
- Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57(4), 460-482.
- Ladefoged, P. (1993). *A Course in Phonetics* (3rd ed.). NY: Harcourt Brace Jovanovich Inc.
- Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *J. Acoust. Soc. Am.*, 29, 98-104.
- Lieberman, A. M., Delattre, P. C., & Cooper, F. S. (1952). The role of selected stimulus variables in the perception of unvoiced stop consonants. *American Journal of Psychology*, 65, 497.
- Magnuson, J. S. (2008). Nondeterminism, pleiotropy, and single word reading: Theoretical and practical concerns. In E. Grigorenko & A. Naples (Eds.), *Single Word Reading*, pp. 377-404. Mahwah, NJ: Erlbaum.
- 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.
- McMurray, B., Tanenhaus, M., & Aslin, R. (2009). Within-category VOT affects recovery from "lexical" garden paths: Evidence against phoneme-level inhibition. *J. of Memory and Language*, 60 (1), 65-91.
- Mirman, D., Dixon, J.A., & Magnuson, J.S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *J. of Memory and Language*, 59, 475-494.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, 24, 175-184.
- Recasens, D. (1984b). V-to-C coarticulation in Catalan VCV sequences: An articulatory and acoustical study. *Journal of Phonetics*, 12, 61-73.
- Recasens, D. & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *J. Acoustical Society of America*, 122(4), 2228-2298.
- Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, 90, 51-89.
- Shillcock, R. C., & Bard, E. G. (1993). Modularity and the processing of closed class words. In G.T.M. Altmann & R. C. Shillcock (Eds.), *Cognitive models of speech processing* (pp. 163-185). Erlbaum.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *J. of Verbal Learning and Verbal Behavior*, 18, 645-659.
- Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken-language comprehension. *Science*, 268, 1632-1634.
- West, P. (2000). Perception of distributed coarticulatory properties of English /l/ and /t/. *Journal of Phonetics*, 27, 405-425.