CAN I [F[W]]*EED* YOU SOME [F[J]]*OOD*?
THE ROLE OF SUBPHONEMIC CUES IN WORD RECOGNITION

Tae-Jin Yoon[1], Anna Moro[1], John Connolly[1], Jessica Arbour[1], and Janice Lam[2]
[1]Dept. of Linguistics and Languages, McMaster University, 1280 Main Street W., Hamilton, Canada, L8S4M2
[2]Dept. of Psychology, Neuroscience, and Behaviour, McMaster University, 1280 Main St. W. Hamilton, Canada, L8S4M2

1. INTRODUCTION

It is well known that a phonological segment often varies to "become more like an adjacent or nearby segment" (Hardcastle & Hewlett 2000) due to overlapping articulatory gestures Coarticulation accounts for the realization of the same phonological segment – for example, the initial /k/ of *key* and *car* – in different areas of the vocal tract. Reserachers agree that coarticulation is the result of the vocal tract producing gestures in 'real time' by transitioning instantaneously from one target configuration to the next. When it comes to the degree, role and function of coarticulation, however, conflicting theories and findings abound. The goal of this study is to find out if and to what extent coarticulatory properties have an impact on spoken word recognition. We designed two experiments: a speech perception experiment, and an ERP (Event Related Potentials) experiment. Both experiments used the same auditory stimuli made with a splicing technique. For example, the initial consonant of *feed* was replaced with the initial consonant of *food*. The spliced stimuli, which contain incorrect sub-phonemic (coarticulatory) properties, were used to investigate the processing of coarticulatory information in word recognition (see Archibald and Joanisse 2011). Thirty-six monolingual English-speaking adults participated in the experiments. We report here preliminary results from both experiments. Results obtained thus far indicate that sub-phonemic cues are indeed utilized by many speakers in spoken word recognition, but the extent to which coarticulatory cues are used depends both on the nature of the initial consonant and on the type of vowel.

2. METHOD

2.1 Stimuli

We used 30 real words of a CVC shape to examine the processing of anticipatory coarticulatory effects on different types of initial consonants. The stimuli contained either an initial stop /p, t, k/ or a fricative /f, s, ʃ, h/. Each monosyllable contained one of the four corner vowels /i, u, æ, ɑ/ for maximal contrast. One of the following consonants filled the coda position: /p, t, d, k, l/. To reduce any possible effects of regressive coarticulation stemming from the final consonant, only minimal sets were used: stimuli sharing an initial consonant also had to contain the same final consonant. Table 1 shows two sets of stimuli used in the experiments, one set displaying an initial stop, and the other an initial fricative. In each set the onset consonant is followed by one of the four contrasting vowels, and the same coda consonant.

Table 1. Examples of stimuli used for splicing

	Stops		Fricatives	
	Front	**Back**	**Front**	**Back**
High	peal	pool	Heat	hoot
Low	pal	Paul	Hat	hot

A female adult speaker of Canadian English produced each word three times. One of the tokens was chosen to prepare the spliced stimulus items. The complete set of stimulus items contained both onsets with congruent coarticulatory acoustic cues (e.g., [f[J]] as the onset of *feed*) and onsets with incongruent coarticulatory information (e.g., [f[w]] as the onset of *feed*).

2.2 Behavioural Speech Discrimination Task

For the behavioural tasks, 20 native speakers of Canadian English were recruited through poster advertisements. Participants were asked to match an auditory stimulus to a pair of stimuli (ABX paradigm).

2.3 ERP Task

For the ERP task, 12 native speakers of Canadian English were recruited through poster advertisements. Electrical activity from the scalp was recorded while the participant listened to auditory stimuli. The written form of the word appeared on a computer monitor as a visual stimulus. Participants were asked to indicate whether the written word and the auditory stimulus matched or not by pressing one of two keys on a handheld keypad.

3. RESULTS

3.1 Behavioural Speech Discrimination Task

Measures of d', β, A', and B''_d were made to analyze the discriminability and bias of the experiment. In this paper, we present the values obtained from A' and B''_d, as in Figure 1. The same interpretation can be made in terms of d' and β, which have correlation values of 0.96 and 0.91, respectively with A', and B''_d. In Figure 1, the mean value of A' is 0.92 and the standard deviation is 0.02. An A' value near 1.0 indicates good discriminability, while a value near 0.5 means chance performance. Participants were able to detect subtle coarticulatory cues. B''_d is a measure of bias, and ranges from -1 to 1 in value. A B''_d value of 0.0 indicates no bias; a number close either to -1 or 1 reflects the degree of bias of a participant (the value calculates how many times the participant clicks either 'yes' alone or 'no' alone without discriminating the correct stimulus). In Figure 1, the mean of B''_d is 0.006, and the standard deviation is 0.372. In general, it appears that most participants can detect coarticulatory cues.

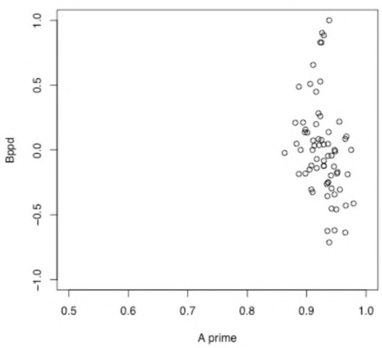

Figure. 1. Scatterplot between the values of A' and the values of B''_d.

The behavioural experiment alone cannot tell us is if participants are merely responding to the presence of deviant acoustic properties or if their behaviour is related to phonological processing.

3.2 ERP studies: the PMN

The most important finding for our purposes is that of the phonological mapping negativity, or PMN (Steinhauer & Connolly 2008). The PMN, a negative-going component (N280) that peaks around the 200-300 ms range, is elicited by a phonological mismatch between the expected and heard onset of a target. The PMN has been understood to be sensitive to phonological processing. More recently, Archibald and Joanisse (2011) investigated whether the N280 is sensitive to a mismatch "between expected [and] perceived phonetic information." Their study provides strong evidence that the PMN is as sensitive to subphonemic information as it is to phonemic information, a finding that argues in favour of the view that subphonemic information is integrated in real time as it unfolds, and aids in word recognition (Gow and McMurray 2007). We designed our stimuli in minimal sets to reduce the effects of coda consonants on the rime, and with a view to controlling for sensitivity to coarticulatory effects based on the onset consonant and the height or backness of the vowel.

As can be seen in Figure 2, which is taken from the central front site in the scalp for the condition of stop consonants, the difference between forms of congruent coarticulatory cues and those of incongruent cues is at its maximum near 300 ms. This confirms that the coarticulatory information is related to phonological expectancy. A further analysis of other conditions indicates that the detection of coarticulatory cues is largely dependent both on the nature of the onset consonant (stop vs. fricative), and on the type of vowel that follows the initial consonant.

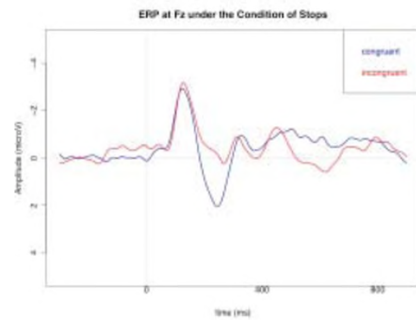

Figure. 2. Average waveform comparing the coarticulatory match and coarticualtory mismatch conditions at the site Fz under the condition of stop consonants

4. DISCUSSION

We have presented two experiments that investigate the processing of coarticulatory information. The speech discrimination experiment showed that speakers are sensitive to coarticulatory cues: many participants detected the coarticulatory mismatch between onset and rime. The neurophysiological study indicates that the detection is not due to pure acoustic deviants, but is done due to its phonological effects. By correlating the behavioral findings with findings from the ERP study, we are aiming to better understand the cognitive and neural processes associated with coarticulatory properties, and the role they play in spoken word recognition.

REFERENCES

Archibald, L. & M. Joanisse. (2011). Electrophysiological Responses to Coarticulatory and Word Level Miscues. *Journal of Experimental Psychology: Human Perception and Performance*, Advance online publication. doi: 10.1037/a0023506.

Steinhauer, K. & J. Connolly (2008) Event-Related Potentials in the study of language. In Stemmer, B. and H. A. Whitaker (eds.) *Handbook of the neuroscience of language*, Academic Press.

Connolly, J. & N. Phillips. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience,* 6: 256-266.

Hardcastle, W. & N. Hewlett (eds.) (1999) *Coarticulation: theory, data, and techniques*. Cambridge: Cambrige University Press.

Gow, J. and B. McMurray (2007). Word recognition and phonology: The case of English coronal place assimilation. In (eds.) J. Cole and J. Hualde, *Papers in Laboratory Phonology 9.* Berlin: Walter De Gruyter: 173-200.

ACKNOWLEDGEMENTS

This work is supported by Arts Research Board (ARB) at McMaster University and NSERC. Statements in this paper reflect the opinions and conclusions of the authors, and are not endorsed by the granting agencies.