PROCESSING DEPENDENCIES BETWEEN SUPRASEGMENTAL AND SEGMENTAL INFORMATION: EFFECTS OF EMOTION AND LEXICAL TONE ON INTERFERENCE

Leher Singh, Quentin Lee & Winston D. Goh

National University of Singapore, Singapore
psyls@nus.edu.sg

ABSTRACT

Processing dependencies in speech have traditionally been studied using a speeded classification paradigm. In this procedure, listeners have to monitor one dimension of speech (e.g. segmental detail) while another dimension (e.g. suprasegmental detail) varies. Variation in the unattended dimension has been shown to interfere with responses to the attended dimension. In the present study, the speeded classification task was adapted to investigate interference from two distinct kinds of suprasegmental information – emotion and lexical tone – in tonal and non-tonal speakers. Results demonstrate a distinct pattern of responses for each group, suggesting influences of native language properties on processing interactions.

Keywords: processing dependencies, indexical variation

1. INTRODUCTION

The speech signal contains multiple tiers of information which include both phonemic detail and non-phonemic detail (e.g. indexical variation). The relationship between these sources of information has traditionally been investigated via a speeded classification paradigm [1]. This paradigm comprises two conditions: control and orthogonal. In the control conditions, listeners are asked to monitor one dimension (e.g. onset consonant) while the other dimension (e.g. talker gender) remains constant. In half of the control trials, the listener monitors phonemic detail while indexical detail remains constant and in the other half, listeners monitor indexical detail (e.g. male or female talker) while segmental information is held constant. Participants' reaction times are then compared with those of an orthogonal condition where listeners again have to either focus selectively on phonemic or non-phonemic detail. Unlike the control conditions, both dimensions now vary. There is typically a cost associated with performance in this condition, as dimensions that listeners are asked to ignore interfere with the attended dimension, resulting in higher response times in orthogonal blocks compared with control blocks [2, 3, 5, 8, 10]. Symmetric interference is observed when i) orthogonal trials are significantly different from control trials for both phonemic and non-phonemic judgments and ii) when the amount of interference is not significantly different for both types of judgments. A significant difference in the amount of interference across both types of judgments reflects asymmetrical interference. In this scenario, one dimension is more closely attended to (and harder to ignore) than the other.

This paradigm has revealed intriguing interactions between phonemic and non-phonemic information processing. For example, talker gender has been found to interfere with processing of onset consonants more so than *vice versa* [3]. In the current study, the relationship between processing dependencies and linguistic function of particular cues is investigated. Processing interactions between lexical tone/consonant changes are investigated in listeners for whom lexical tone is phonemic as well as those for whom it is not. In addition, emotion/consonant change interactions are investigated in the same cohort of participants. Emotion shares many of the acoustic drivers of lexical tone yet is considered a source of non-phonemic variation for both groups, although emotional prosody does constrain semantic interpretations of speech [4].

Our hypotheses were that 1) Chinese speakers would show symmetrical interference with lexical tone/consonant variation (i.e. equal interference from tone variation on consonant judgments as from consonant variation on tone judgments) given the phonemic status of both lexical tone and the target consonants in their language [2, 5]) 2) Chinese speakers would also show symmetrical interference in emotion/consonant conditions due to the necessity of integrating emotional prosody and consonant information into semantic analysis 3) English listeners would show asymmetric interference with Mandarin tones showing greater interference from consonants on tones than *vice versa* and 4) English listeners would show symmetrical interference between emotional changes and consonant changes for the same reasons as those articulated for Mandarin listeners.

2. METHODS

2.1. Participants

Sixty-four participants were recruited for the study. Thirty-two participants were native speakers of Mandarin Chinese and 32 were native speakers of English with no tone language experience.

2.2. Stimuli

All stimuli consisted of natural productions of /ba/ and /da/ recorded by a Chinese speaker. Tone stimuli consisted of Tones 2 (rising) and 4 (falling), chosen for their relatively large tonal range, distinct tonal contours and that they are relatively well discriminated by non-tonal speakers [9]. Emotion stimuli consisted of natural productions of happy and sad instantiations of /ba/ and /da/ and were recorded by the same speaker. Stimulus validation procedures, performed by Chinese-English bilinguals, revealed near-perfect classification of tone stimuli (95%) and emotion stimuli (98%) (n=10). Emotional intensity ratings were 4.87 on a 5 point scale for happy stimuli and 1.27 for sad stimuli.

2.3. Procedure

The experiment comprised two within-subject conditions. One condition (emotion/consonant condition) consisted of 6 blocks of 40 trials each with emotion as the suprasegmental variable and onset consonant identity as the segmental variable. The other condition (tone/consonant condition) consisted of 6 blocks of 40 trials each with lexical tone as the suprasegmental variable and onset consonant identity as the segmental variable. For the purposes of simplicity and consistency with previous studies, lexical tone will be referred to as the suprasegmental dimension although it is clearly segmental for Mandarin Chinese listeners. In each condition, there were two paired control blocks and one orthogonal block. In the control blocks, the attended dimension (e.g. consonant) varied while the unattended dimension (e.g. emotion) was held constant. In the orthogonal blocks, both dimensions varied. All blocks were presented twice and in one presentation participants were asked to make segmental judgments (classify syllables as 'ba' or 'da') while in the other, they were asked to make suprasegmental judgments (classify syllables as 'happy' or 'sad' for the emotion condition or 'tone 2' or 'tone 4' in the tone condition. The tonal system was explained to English speakers prior to the test and examples were provided of each tone. Stimuli were presented over headphones. Response times greater than 5000 msec or less than 300 msec were excluded (leading to exclusion of 2% of trials for each group). Reaction times were measured from stimulus onset. The experimental design is displayed below.

Table 1a: Experimental design for emotion condition.

<table>
<tr><th>Target Dimension</th><th>Control</th><th>Orthogonal</th></tr>
<tr><td>Segmental</td><td>Block 1:
/ba/ - Happy
/da/ - Happy
Block 2:
/ba/ - Sad
/da/ - Sad</td><td rowspan="2">/ba/ - Happy
/da/ - Sad
/ba/ - Happy
/da/ - Sad</td></tr>
<tr><td>Suprasegmental</td><td>Block 1:
/ba/ - Happy
/ba/ - Sad
Block 2:
/da/ - Happy
/da/ - Sad</td></tr>
</table>

Table 1b: Experimental design for tone condition.

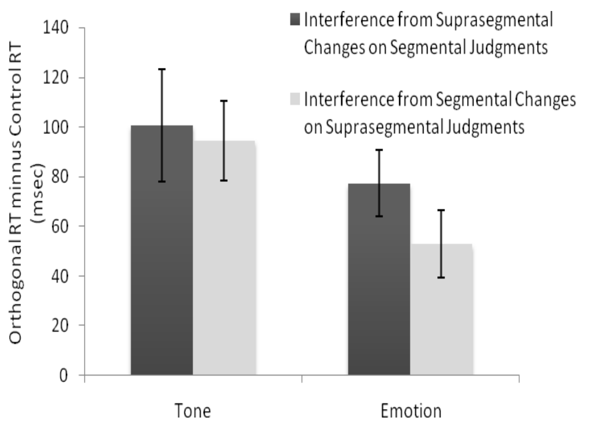
<table>
<tr><th>Target Dimension</th><th>Control</th><th>Orthogonal</th></tr>
<tr><td>Segmental</td><td>Block 1:
/ba/ - Tone 2
/da/ - Tone 2
Block 2:
/ba/ - Tone 4
/da/ - Tone 4</td><td rowspan="2">/ba/ - Tone 2
/da/ - Tone 2
/ba/ - Tone 4
/da/ - Tone 4</td></tr>
<tr><td>Suprasegmental</td><td>Block 1:
/ba/ - Tone 2
/ba/ - Tone 4
Block 2:
/da/ - Tone 2
/da/ - Tone 4</td></tr>
</table>

3. RESULTS

As in previous studies, processing dependencies were established by comparing performance in control and orthogonal blocks for segmental and suprasegmental judgments. Across all tasks, accuracy rates did not differ for tone and emotion conditions for either group of speakers (range: 94 to 98%). Statistical analyses therefore focused on reaction times as in previous studies [2, 3, 5]. A preliminary set of analyses was conducted to determine i) whether there was evidence of interference in each condition and ii) whether baseline performance (control block RTs) differed for segmental and suprasegmental dimensions. A 2 X 2 ANOVA (control/orthogonal x segmental/suprasegmental) was conducted to determine whether there was evidence of interference in each condition. There was a significant main effect of block type (control/orthogonal) for both groups of speakers on both tone and emotion tasks, (Chinese: F (1,62) = 55.37, p<.001 (tone) and F (1,62) = 38.61, p<.0001 (emotion); English: F (1,62) = 44.38, p<.0001 (tone) and F (1,62) = 59.47, p<.0001 (emotion)). A oneway ANOVA yielded no significant differences in control RTs in suprasegmental or segmental judgments (Chinese: F (3,124) = .11, NS, English: F (3,124 = .39, NS)).

To investigate our hypotheses, the amount of interference (response times to orthogonal blocks minus response times to control blocks) was entered into a 2 X 2 ANOVA for each language group with condition (tone/emotion) and judgment type (segmental/suprasegmental) as factors. Results for the Chinese group are displayed in Figure 1a.

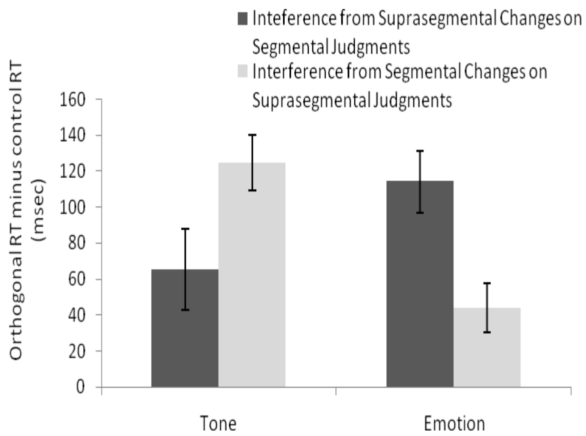
Figure 1a: Interference effects for Chinese speakers for emotion and tone tasks (errors bars: SEM).



Results revealed no main effect of judgment type (segmental or suprasegmental), no interaction of condition (tone/emotion) and judgment type, but a main effect of condition, F (1,62) = 4.1, p<.05 whereby interference effects were greater for the tone condition than the emotion condition. This pattern of results reveals mutual and symmetrical interference between suprasegmental and segmental judgments whether the suprasegmental manipulation is tone or emotion.

English speakers showed a different pattern of results (see Figure 1b): no significant main effect of judgment type (segmental/suprasegmental), no significant main effect of condition (tone/emotion), but a significant cross-over interaction of judgment type and condition, F(1,62) = 16.22, p<.0001. Simple effects tests revealed significantly greater interference from segmental variation when making tonal judgments than from tonal variation when making segmental judgments, t(31) = 2.2, p<.05. By contrast, in the emotion task, English speakers showed significantly more interference from emotion variation when making segmental judgments than *vice versa*, t(31) = -3.96, p<.0001.

Figure 1b: Interference effects for English speakers for emotion and tone tasks (error bars: SEM).



4. DISCUSSION

There was evidence of interference from the unattended dimension on the target dimension in each condition for both language groups. However, there were marked differences in the amount of interference across conditions and language groups. Hypothesis 1 (H1) stated that Chinese speakers would show symmetrical interference with lexical tone/consonants. This hypothesis was confirmed, adding to previous evidence of symmetrical tone/consonant interference in Chinese listeners in studies using synthesized speech [2, 5], an effect attributed to the phonemic status of both lexical tone and consonant identity in Chinese. Hypothesis 2 (H2) stated that Chinese speakers would also show symmetrical interference in the emotion/consonant condition. This too was confirmed. Evidence of symmetrical interference for both tone/consonant and emotion/consonant judgments may be attributable to the lexical and communicative functions served by tones, emotional prosody and consonant identity for Mandarin Chinese listeners. Therefore, all three sources of information may be processed with comparable priority and may be equally difficult to disregard.

Hypothesis 3 (H3) stated that English listeners would show asymmetric interference with Mandarin tones and consonant changes. This hypothesis was confirmed: English speakers demonstrated less interference from the non-phonemic suprasegmental variation (tone) on phonemic segmental judgments (consonant) than *vice versa*, suggesting that phonemic detail is less vulnerable to interference than non-phonemic suprasegmental detail. Thus far, results demonstrate symmetrical interference when both dimensions are functionally relevant (H1 and H2) and processing asymmetries when only one dimension is functionally relevant (H3).

Hypothesis 4 stated that English listeners would show symmetrical interference between emotional changes and consonant changes due to the communicative relevance of emotional prosody. This hypothesis was not upheld and English listeners showed more interference from suprasegmental emotional detail on segmental judgments than *vice versa*. It should be noted that English speakers showed marginally greater interference from emotion on consonant judgments than did Chinese speakers ($t(31) = 1.8$, $p=.1$). While this finding accords with the findings of Mullenix and Pisoni [3] on talker gender, it is not clear why English and Chinese speakers would demonstrate a different pattern of results in the emotion condition given that prosodic cues to emotion are relevant to English and Chinese speakers presumably in equal measure. One possibility could be that differences exist in prosodic cues to emotion recruited in tone and non-tone languages. Tone languages may rely less on prosodic cues to communicate emotion possibly due to the potential lexical ambiguity that may ensue [6]. Therefore, Chinese speakers may be less sensitive to prosodic cues to emotion than English speakers. This hypothesis requires empirical scrutiny as there have been no studies thus far contrasting perception of emotional prosody in tonal and non-tonal language speakers.

In summary, results support a language-specific account of processing dependencies. While both language groups demonstrated interference from unattended dimensions, the degree of symmetry of interference appeared to be influenced by the contrastive phonology of listeners' native languages.

5. REFERENCES

[1] Garner, W. 1974. *The Processing of Information and Structure.* Potomac, MD: Erlbaum.
[2] Lee, L., Nusbaum, H. 1993. Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception and Psychophysics* 53, 157-165.
[3] Mullenix, J., Pisoni, D. 1990. Stimulus variability and processing dependencies in speech perception. *Perception and Psychophysics* 47, 379-390.
[4] Nygaard, L., Queen, J. 2008. Communication emotion: Linking affective prosody and word meaning. *Journal of Experimental Psychology* 34, 1017-1030.
[5] Repp, B., Lin, H. 1990. Integration of segmental and tonal information in speech perception: A cross-linguistic study. *Journal of Phonetics* 18, 481-495.
[6] Ross, E., Edmondson, J., Seibert, G. 1986. The effect of affect on various acoustic measures of prosody in tone and non-tone languages. A comparison based on computer analysis of voice. *Journal of Phonetics* 14, 283-302.
[7] So, C., Best, C. 2010. Cross-language perception of non-native tonal contrasts: Effect of native phonological and phonetic influences. *Language and Speech* 53, 273-293.
[8] Tong, Y., Francis, A., Gandour, J. 2008. Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes* 23, 689-708.
[9] Wang, Y., Spence, M., Jongman, A., Sereno, J. 1999. Training American listeners to perceive Mandarin tones. *The J. Acoust. Soc. Am.* 106, 3649-3658.
[10] Wood, C. 1974. Parallel processing of auditory and phonetic information in speech discrimination. *Perception and Psychophysics* 15, 501-508.