DIFFERENTIAL USE OF TONAL AND SEGMENTAL INFORMATION IN LEXICAL DECODING DECISIONS IN CANTONESE

Rosemary Varley, Sandra P. Whiteside & Yuet Yee Yim
Speech Science, University of Sheffield, Sheffield S10 2TA, United Kingdom

ABSTRACT

The study investigated the perceptual processing of tonal and segmental information in Cantonese. The hypothesis that the processing of tonal information was more robust was tested by presenting 40 young adult subjects with stimuli masked by white noise. Two experimental conditions were developed: free-choice and forced-choice. In both conditions, the results showed a significant primacy of tonal information over segmental information.

INTRODUCTION

The project investigated whether speakers of a tone language (Cantonese) show a preferential use of tonal over segmental information in lexical decoding. Acquisitional research supports the early acquisition of tonal information over segmental [1], and anecdotal evidence from L2 learners of Cantonese suggests that whereas segmental errors can be assimilated by native Cantonese speakers, tonal errors have more profound effects on comprehension. Specifically, where the L2 learner makes a tonal error, the listener appears to make an assumption that the tonal information is as given, but that the segmental form is at fault. Evidence from brain-damaged speakers also suggests that segmental information is more vulnerable to disruption than tonal [2].

These observations suggest a primacy in the processing of tonal information over segmental. This study subjected this hypothesis to experimental test by placing listeners in marginal listening conditions and observing whether there was preferential use of tonal or segmental information.

The tonal system of Cantonese consists of six contrastive tones [3]. There are three level tones, differing in pitch height; high-level (tone 1), mid-level (tone 3), and low-level (tone 6), and three contour tones; high-rise (tone 2), low-rise (tone 5), and low-fall (tone 4). There are also three clipped or entering tones, but these are regarded as allophonic variants of the high, mid, and level tones as they occur only in CVC syllables where the final consonant is /p, t, k/. Tones which share similar contours or starting heights have been shown to be confusable [3, 4, 5].

METHOD

In the first stage of the study, 47 triplets of words were identified. Each triplet included a stimulus word, a tonal minimal pair (same segmental form as the target, but different tone) and a segmental minimal pair (same tone, different segmental form). For example, for target /kun3/, tonal pair /kun1/, segmental pair /pun3/. The minimal pairs were all highly confusable with the target item. The tonal pairs involved tones 1-3, 2-4, 3-4, and 2-6. (There is also a further sub-set of highly confusable tone pairs which include tones 4-6, 3-6, and 2-5, but we were unable to identify triplets including these pairs). Segmental confusability norms are not available for Cantonese and so extrapolations were made from English norms [6]. Segmental pairs involved the following contrasts: /p-k, t-p, ph-th, ph-kh, th-kh, s-f, j-w, j-l, w-l/.

Stimulus Preparation

An adult female native speaker of Cantonese (PY) read the list of target stimuli in a sound proof room. This list also included an introductory phrase (IP) and a prompting phrase (PP). The recording was made on a SONY TCD-D3 DAT recorder. The word stimuli, IP, and PP were digitised using a KAY Computerised Speech Lab (CSL) model 4300 at a sampling rate of 20 kHz. The amplitudes of the word stimuli were then scaled to a mean of 53 dB using the CSL.

These lists were used to record the test stimuli with white noise from the CSL onto a SONY DAT machine. The

white noise was sampled at 20kHz and had an amplitude of 62 dB, to give a mean signal to noise ratio value of -9dB. The speech signal was output to the left channel and the white noise to the right channel. The two signals were then output through a mono channel which was then converted into a stereo channel so that both the white noise and the speech stimuli would be played back in both ears. This was done to control for any bias effects which could have resulted from selective attention listening strategies.

Each set of stimuli included two practice items. An IP was included at the start of both the practice items and the listening items. Each stimulus was preceded either by an IP (at the start) or a PP. This was done to draw the attention of the listener to the stimuli and to familiarise the listeners with the speaker's voice. This is particularly important in tonal perception, where a decision regarding the identity of a tone is relative to the speaker's indexical pitch.

The stimuli were arranged into four random orders. As subjects completed two experimental conditions, each subject received the stimuli in different random orders across conditions, and within conditions, half of the subjects received stimuli in one random order, and the other half in a second.

Subjects

Forty young adult subjects (20 male, 20 female) completed the task. Ages ranged from 19 to 26 (mean 21.8, *sd.* 1.68). All subjects had Cantonese as their L1 and their home language. They reported no hearing problems. All subjects were students in tertiary education.

Procedure

Two experimental designs were developed to test the hypothesis: free-choice and forced-choice. In the free-choice design subjects were required to listen to a stimulus word and to write down what they thought they had heard. Responses were examined for the relationship between the stimulus and the character written down and were analysed into the following categories: stimulus; tonal-response (same tone, different segmental form); segmental response (same segmental form, different tone); and 'other' (for example, both the tone and the segmental form differed from that of the stimulus). In the forced-choice design, subjects were played a distorted syllable and were required to select from two choices the syllable that they thought matched the one they had heard. The two response choices were the segmental minimal pair and the tonal minimal pair, but the actual stimulus word was not given as a response choice. Preferential use of tonal or segmental information was noted.

All subjects were tested individually in sound-damped booths. Test tapes were played on a Sony Stereo Cassette-Corder (TC-D5M), and subjects listened through AKG Dynamic System (K135) headphones. The test procedure took approximately 30 minutes for each subject. The order of the free and forced-choice tasks was not counterbalanced. The free-choice task was presented first so that given responses for the forced-choice task did not influence response in the free-choice condition.

RESULTS

Due to occasional errors in the data collection phase, not all subjects made decisions on all 47 stimuli. Data were therefore converted into percentage scores to permit comparisons.

Free-choice condition

Table 1 shows the percentages of responses classified into each category.

Table 1. Percentage of free-choice responses by category and sex.

	Stim.	Tonal	Segm.	Other
Male				
mean	41.48	42.98	6.52	8.97
sd.	7.74	7.56	4.07	3.20
Fem.				
mean	42.20	42.87	5.39	9.51
sd.	10.12	6.56	3.21	6.91

The pattern of performance between male and female subjects was very similar and therefore data from the two groups were combined for subsequent analysis. The data show that despite the distortion of the stimulus word, subjects were able to decode the tonal information of the syllable on approximately 84 percent of trials. Segmental information was more vulnerable to disruption and was accurately decoded on approximately 47 percent of trials. It was rare for segmental information to be decoded but for the tone not to be correctly perceived.

Differences between response categories were compared with Wilcoxon signed-rank tests. Comparisons were non-significant between target and tonal responses. All other comparisons were significant (target score > segmental and other scores $p>0.01$ ($T = 0$ and 1 respectively), tonal scores > segmental and other scores $p>0.01$ ($T = 0$), and other score >segmental score $p>0.01$ ($T = 160$).

Forced-choice condition

Table 2 shows the percentages of tonal or segmental responses made by subjects on the forced choice condition.

Table 2. Percentage of tonal and segmental responses on the forced-choice task, by sex.

		Tonal	Seg.tal
Male	mean	62.65	37.35
	sd.	21.76	
Female	mean	68.26	31.75
	sd.	20.74	

Inspection of the data suggests that again there is no obvious difference between male and female subjects and therefore their results were combined. Subjects preferentially used tonal information when placed in the forced-choice situation. Comparison with a Wilcoxon-signed ranks test revealed the difference to be significant ($T = 126.5, p<0.01$). It is noticeable, however, that subjects utilised a segmental-decoding strategy more often than one might predict from the free-choice situation. Certain individuals showed a marked preference for a particular response type. For example, 13 subjects chose the tonal response on over 80% of trials; strong segmental response choices were less evident, with only two subjects producing 70% or more segmental responses. Spearman correlation coefficients were calculated between tonal responses on the free and forced choice paradigms, and also on segmental responses across the two paradigms. Both figures for *rho* were very low and non-significant (tonal choice *rho* = 0.196, segmental *rho* = 0.001). The higher segmental responses may therefore reflect the effects of chance, and also in the more constrained forced-choice task, the results of selective attention strategies.

DISCUSSION

The results of the perceptual experiments revealed findings consistent with the initial hypothesis that the decoding of tonal information showed primacy over segmental decoding. Subjects were able to perceive the correct tonal information on over 80 percent of trials on a free-choice task, whilst the segmental information was correct on less than half of the trials. Similarly, when placed in a situation where the listener had to choose a character which matched a distorted stimulus, subjects showed a preference for selecting the response choice which matched the target in tone, but not in segmental form.

There are a number of possibilities in accounting for this result. The first is, in line with our experimental hypothesis, that the perception of tone is in some way primary in the recognition systems of tone language users. But primacy here does not mean temporal primacy. Segmental and tonal perceptual processes are likely to happen in parallel: one could not envisage a perceptual system which 'held' on segmental information until a tonal decision had been made. It may be that the processing of tonal information is a more rapid process. On a mathematical basis, there are only six contrastive tones in Cantonese, but many more possible segments and their combinations. As the tone paradigm is smaller, then decisions will be made more quickly. This might then suggest that the phonological lexicon will be organised around tone - the rapid tonal decision permitting a narrowing down of the possibilities in matching an input to a stored representation.

A second possibility is that our results are more simply an artefact of our experimental design. There are two possibilities here. The first is that our segmental minimal pairs were more confusable than the tonal ones. Whereas it was possible to select or extrapolate from English norms highly confusable consonants, we were unable to identify triplets including the very highly confusable tone pairs. The absence of such triplets may be an example of avoidance of contrasts in a single language which pose heavy demands on perceptual systems.

A second explanation of the results is that the white noise masking is more likely to distort the acoustic information necessary for segmental (especially consonantal) perception than the lower frequency information involved in tonal discrimination. Preliminary analysis of free-choice errors supported the idea that some errors were an artefact of the experimental design. Consonant errors for example were more frequent than vowel errors: suggesting that the lower frequency information of vowels and diphthongs (and also therefore tones) was more robust in the white noise masking used in this study. It was felt however that all the results could not be totally dismissed as artefacts. The experiment used blanket masking of the signal, but the speech signals varied temporally in amplitude. This meant that the vowel/ diphthong nuclei and formant transitions in particular were most robust to the masking than consonantal features like frication and plosion. Given this we would have expected listeners to use the formant transition patterns to opt for more segmental judgements over tonal judgements. However as was noted above this was not the case and listeners tended to opt for tonal judgements. Furthermore some of the distinctions between the tone pairs used in this study were signalled at the tail-end of the vowel [3], and therefore were subject to the same level of masking as consonant cueing transitional information. Despite this, listeners tended to opt for the tonal decisions. A study on Mandarin Chinese tones [7] found that the perception of tone was robust to various filter conditions. Here listeners were still able to make tone-phoneme identifications with missing acoustic information. This lends further support to our findings which illustrate the primacy of tone in the perceptual systems of tone language users. Further research is planned in this area.

ACKNOWLEDGEMENTS

The authors would like to thank the Department of Speech and Hearing Sciences at the University of Hong Kong for use of their facilities during the data collection phase of this project, and P. Yip for her help in recording the stimuli.

REFERENCES

[1] Tse, J.K.P. (1978), Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language* vol. 5, pp. 191-204.

[2] Yiu, E. M. L. (1989), *Tonal Disruption in Chinese (Cantonese) Aphasics*. Unpublished M. Phil. Thesis. University of Hong Kong.

[3] Fok, C.Y.Y. (1974), *A Perceptual Study of Tones in Cantonese*. Hong Kong: University of Hong Kong.

[4] Gandour, J. (1983), Tone perception in Far Eastern languages. *Journal of Phonetics* vol. 11, pp. 149-175.

[5] Varley, R. & So, L. (1995), Age effects in tonal comprehension in Cantonese. In press. *Journal of Chinese Linguistics*.

[6] Miller, G. & Nicely, P. (1955), An analysis of perceptual confusion among some English consonants. *Journal of the Acoustical Society of America*, vol. 27, pp. 338-352.

[7[Stagray, J. R. , Downs, D. & Sommers, R. K. (1992), Contributions of the Fundamental, resolved harmonics, and unresolved harmonics in tone-phoneme identification. *Journal of Speech and Hearing Research,* vol. 35, pp. 1406-1409.