Mandarin Chinese Speech Perception in Noise: Phonological Implications

Kevin Tang[1], Yan Lou[2]
University College London, Department of Linguistics
[1]kevin.tang.10@ucl.ac.uk, [2]lslylenalgg@gmail.com

Keywords: Speech Perception, Tone, Perceptual Similarity

When mentioning similarity of speech sounds, one must consider the two classical large scale confusion experiments by Miller & Nicely (1955) and Wang & Bilger (1973) on English. Very few studies worked on non-English confusion, e.g. Singh & Black (1966) on Hindi, Arabic and Japanese. The focus of these studies also tended to be rather narrow (consonants, or vowels) or use only nonsense syllables (CVC). Following the footsteps of Cutler et al. (2004), we present a large scale confusion study of Mandarin Chinese for consonants, vowels and tones with all attested CV and VC syllables, with a focus on its implications in phonology from the confusion patterns of consonants, vowels and, most interestingly, tones.

Methods: All attested CV (704) and VC (12) syllables with attested tones were embedded in speech-shaped noise with three levels of SNRs (at -8 dB, 0 dB and +8 dB). Thirty Mandarin native listeners were tested with an "open-set response" method.

Key Findings: Tone 2 and Tone 3 were found to be the most confusable pair (See Table 1). This can be accounted for with the well-known tone sandhi of Tone 3, which creates ambiguities in the tone sequence Tone 2 - Tone 3 having the same surface form as the tone sequence Tone 3 - Tone 3 (Duanmu 2002). Hume and Johnson (2003) suggested that since these two tones undergo contextual neutralization, they therefore carry

Tone Confusion	Percentage	[Contour] Change
Tone 2 – Tone 3	3.69%	1
Tone 3 – Tone 4	2.15%	1
Tone 1 – Tone 2	1.81%	1
Tone 2 – Tone 4	0.67%	0
Tone 1 – Tone 4	0.59%	1
Tone 1 – Tone 3	0.49%	0

Table 1: Tone confusion pairs

less functional load than contrastive pairs, which in turn causes them to be partially contrastive. In fact, this tonal-pair was calculated to have the lowest functional load amongst the four tones (Surendran & Levow 2004). Furthermore, studies in child tone acquisition confirmed our findings; Tone 3 tends to perform worst among the four tones (Wong, Schwartz, & Jenkins 2005), and Tone 2 and Tone 3 not only acquired after Tone 1 and Tone 4, but are relatively confusable throughout the whole process of tone acquisition (Li & Thompson 2008).

By considering confusion as transmitted information, a sequential information analysis (SINFA) (Wang & Bilger 1973) was performed. Using a two feature system [Contour] and [StartHigh] for Mandarin tones (Yip 2002), in which Tone 2 and Tone 4 have [+Contour], Tone 1 and Tone 4 have [+StartHigh], SINFA showed that information transmitted of [Contour] is 86.1% and that of [StartHigh] is 90.02%, suggesting that the feature [Contour] is less robust than [StartHigh]. This is reflected in the confusability of tone pairs, the tone pairs that have a feature change in [Contour] have relatively higher confusion rate, with the exception of Tone 1-4 pair (Table 1). The poor robustness of [Contour] can be phonetically explained, since contour tones demand sufficient duration in perceptual decoding of pitch change (Zhang 2004).

The three palatal fricatives in Mandarin [tɕ, tɕʰ, ɕ] are in complementary distribution with three other sets of sounds, the velars [k, kʰ, x], the dentals [ts, tsʰ, s], and the retroflexes [tʂ, tʂʰ, ʂ]. The underlying representations of these palatals have been a long-debated topic; Chao (1934) proposed the velars as the UR, Hartman (1944) and Duanmu (2007) argued for the dentals and Cheng (1973) argued for their independence. Our results showed that the palatals [tɕ, tɕʰ, ɕ] were most confusable with the retroflexes [tʂ, tʂʰ, ʂ], therefore they are perceptually more similar. This suggests that at least from the angle of phonetic similarity of allophones, the retroflexes are the most likely UR candidates. This agrees with Lu (2011)'s priming study showing that the

relationship between [s] and [ɕ] is more phonemic than allophonic.

Finally, bilabial and velar plosives were distinctively more confusable with each other – [pʷ] > [kʷ] (17.4%), [pʰʷ] > [kʷ] (6.8%), [pʰʷ] > [kʰʷ] (5.1%), [kʷ] > [pʷ] (3.8%) and [kʰ] > [pʰ] (3.6%) (">" denotes the direction of confusion). These confusions provided further evidence for the feature [grave], since the feature [+grave] groups some labial and dorsal consonants in terms of their spectral properties (Ladefoged 2011; Backley & Nasukawa 2009).

Our work reinforces the complementarity of naturalistic, experimental, and phonological analyses. The findings allowed us to better understand the segmental and suprasegmental phonology of Mandarin and the language-particular effects of speech perception in noise.

References

Backley, P., & Nasukawa, K. (2009). Representing Labials and Velars: A Single 'Dark' Element. *Phonological Studies, 12,* 3-10.

Chao, Y. R. (1934). The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology, Academia Sinica* IV, 363-397.

Cheng, C. C. (1973). *A synchronic phonology of Mandarin Chinese* (Vol. 4). Walter de Gruyter.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116,* 3668.

Duanmu, S.. (2007). *The Phonology of Standard Chinese*. Oxford University Press.

Hartman, L. M. (1944). The segmental phonemes of the Peiping dialect. *Language* 20, 28-42.

Hume, E., & Johnson, K. (2003). The impact of partial phonological contrast on speech perception. In *Proceedings of the XVth International Congress of Phonetic Sciences*.

Ladefoged, P. (2011). Hierarchical Features of the International Phonetic Alphabet. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (Vol. 14).

Li, C. N., & Thompson, S. A. (2008). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, *4*(02).

Lu, Y. A. (2011). The psychological reality of phonological representations: the case of Mandarin fricatives. In *23rd North American Conference on Chinese Linguistics (NACCL), June* , 17-19.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, *27,* 338.

Singh, S., & Black, J. W. (1966). Study of Twenty-Six Intervocalic Consonants as Spoken and Recognized by Four Language Groups. *The Journal of the Acoustical Society of America*, *39,*372.

Surendran, D., & Levow, G. A. (2004). The functional load of tone in Mandarin is as high as that of vowels. In *Speech Prosody 2004, International Conference*.

Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. *The Journal of the Acoustical Society of America*, *54*(5), 1248–1266.

Wong, P., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of speech, language, and hearing research : JSLHR*, *48*(5), 1065–1079.

Yip, M. (2002). Tone. Cambridge University Press.

Zhang, J. (2004). Contour tone licensing and contour tone representation. *Language and Linguistics*, *5*(4), 925-968.