

Gradient phonemic contrast in Nanjing Mandarin

Keith Johnson (UC Berkeley)

Yidan Song (Nanjing Normal University)

Abstract

Sounds that are contrastive in a language are rated by listeners as being more different from each other than sounds that don't occur in the language or sounds that are allophones of a single phoneme. The study reported in this paper replicates this finding and adds new data on the perceptual impact of learning a language with a new contrast. Two groups of speakers of the Nanjing dialect of Mandarin Chinese were tested. One group was older and had not been required to learn standard Mandarin as school children, while the other younger group had learned standard Mandarin in school. Nanjing dialect does not contrast [n] and [l], while standard Mandarin does. Listeners rated the similarity of naturally produced non-words presented in pairs, where the only difference between the tokens was the medial consonant. Pairs contrasting [n] and [l] were rated by older Nanjing speakers as if two [n] tokens or two [l] tokens had been presented, while these same pairs were rated by younger Nanjing speakers as noticeably different but not as different as pairs that contrast in their native language.

Introduction

The variety of Mandarin Chinese spoken in Nanjing does not have a contrast between [n] and [l] (Song, 2015), while the standard variety which is now taught in schools in Nanjing does have this contrast. Prior research has shown that speech perception is modulated by linguistic experience. Years of research has shown that listeners find it very difficult to perceive differences between sounds that are not contrastive in their language (Goto, 1971; Werker & Tees, 1984; Flege, 1995; Best et al., 2001). More recently the role of allophony in perception has been studied, finding that sounds in an allophonic relationship in a language are not as perceptually distinct as sounds that are contrastive (Boomershine et al., 2008; Huang & Johnson, 2010; Johnson & Babel, 2010).

The present study concerns the impact of another type of linguistic experience. This experiment tested whether second language acquisition may influence sound similarity. This was done by comparing older Nanjing dialect speakers, who have limited experience with standard Mandarin with younger speakers of the Nanjing dialect who learned standard Mandarin from an early age in school.

The experiment reported here was designed to test the subjective perceptual consequences of lack of [n]/[l] contrast in Nanjing dialect, and whether learning the standard variety of Mandarin in school would result in an increase in sensitivity to the contrast. We compared the performance of the Nanjing listeners with a group of listeners whose native language contrasts [n] and [l], namely American English. Tables 1 and 2 show consonant inventories of English and Nanjing Mandarin. Note that the /z/ phoneme in Nanjing dialect can be pronounced [ɹ] (Chao, 1929). Recent evidence (Liao & Shi, 1987; and Lin, 1995) shows that the Nanjing retroflex voiced fricative /z/ is pronounced without frication noise [ɹ] intervocalically, especially before [a], and [z] at word onset.

Methods

Listeners. Thirty-two speakers of Nanjing dialect participated in this study. They were recruited and the experiment conducted at Nanjing Normal University. One group of 16 Nanjing speakers (9 women, 7 men) was composed relatively young people (mean age = 30.9, min = 19.9, max = 45.6), while the other group of 16 (8 men, 8 women) were older (mean age = 61.4, min = 48.6, max = 74.6). In addition, we ran a group of 35 college students at UC Berkeley.

Table 1. English consonants

Labial	Dental	Alveolar	Palato-alveolar	Velar
p b		t d		k g
			tʃ ʤ	
f v	θ ð	s z	ʃ ʒ	
m		n		ŋ
		l		
		ɹ		

Table 2. Nanjing consonants (Song, 2015)

Labial	Alveolar	Retroflex	Alveolo-Palatal	Velar
p ^h p	t ^h t			k ^h k
	ts ^h ts	tʂ ^h tʂ	tɕ ^h tɕ	
f	s	ʂ z _l =ɹ	ɕ	x
m				ŋ
	l~n			

Stimuli. A male native speaker of Nanjing Mandarin pronounced 7 instances each of the 3 pseudowords [tana], [tala] and [tara]. These tokens were all produced with falling tone (Nanjing dialect, tone 1) on the first syllable and a low rising tone (Nanjing dialect, tone 2) on the final syllable. This speaker can also speak standard Mandarin and so could produce the contrast between [l] and [n]. Figure 1 shows spectrograms, waveforms, and pitch traces of three of the stimuli.

The range of acoustic variation among the different tokens of the pseudowords was judged auditorily to be remarkably small, and the tone and duration patterns for the different words was likewise extremely similar. Therefore all 7 instances of each pseudoword was used in the experiment. The tokens were arranged in pairs for the experiment by fully crossing them for a total of $441 = (7 \times 3)^2$ pairs. Thus one third of the pairs were composed of tokens of the same pseudoword (e.g. [tana]/[tana]), while the remaining two thirds were composed of different pseudowords (e.g. [tana]/[tala]).

Task. Tokens were presented in pairs diotically (identical signal to both ears) over earphones at a comfortable listening level. The tokens in a pair were separated by an inter-stimulus interval of 1 second. Listeners were asked to rate the subjective difference between tokens in each pair using a Likert scale where 1 = similar and 5 = different. They indicated their response by pressing a number on a computer keyboard. The program waited for the listener's response, however long it took. If the listener pressed a number not in the range from 1 to 5 a message was displayed asking them to restrict

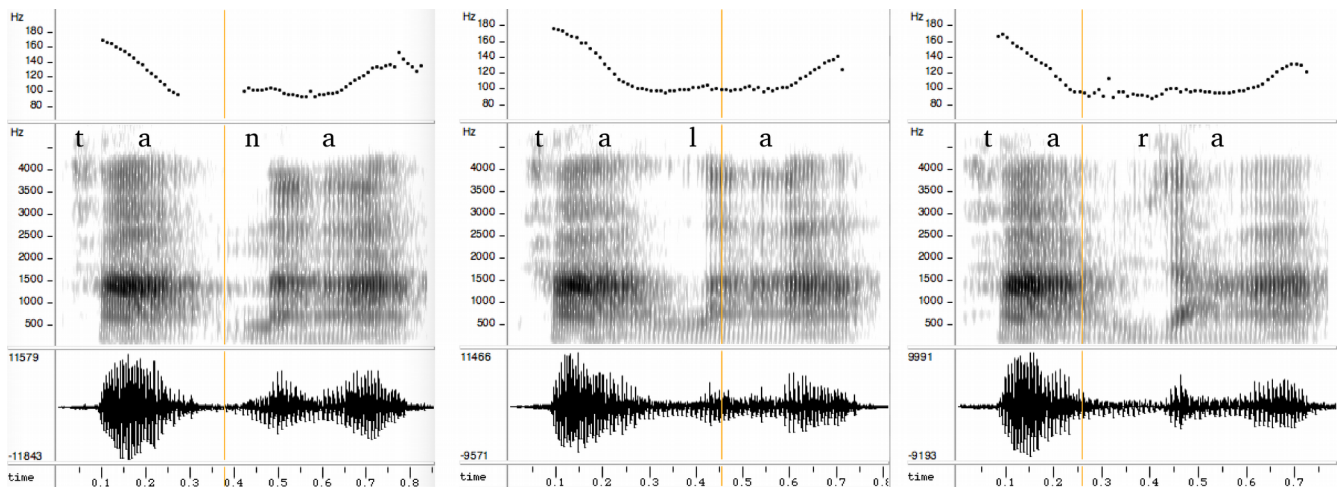


Figure 1. Spectrograms and pitch traces of three of the stimuli – [tana], [tala], and [tara]. Text labels are placed over the acoustic midpoint of each segment. The final vowel in these utterances has a low rising tone and a bi-sonorous pattern – with two peaks of amplitude.

their responses to this range. After the response the listener saw a feedback screen for 2 seconds. The feedback screen showed the current trial number (out of 441), the response that was just given, and the response time. There was no instruction to respond as quickly as possible, but we did want to give participants some motivation to respond with their first impression. Listeners were given breaks after trials 147 and 296.

Table 1. Average ratings (1 = similar sounding; 5 = different sounding)

pair	American English	Nanjing (young)	Nanjing (old)
[l]/[n]	3.34	2.29	1.53
[n]/[l]	3.02	2.13	1.41
[l]/[r]	3.77	4.71	4.81
[r]/[l]	3.79	4.54	4.74
[n]/[r]	3.82	4.74	4.83
[r]/[n]	3.92	4.57	4.82
[l]/[l]	1.51	1.53	1.28
[r]/[r]	1.58	1.44	1.40
[n]/[n]	1.90	1.57	1.35

Results

The average rating (Table 1) depended to minor degree on the order of presentation in five of the nine comparisons for trials where different pseudowords were presented (3 groups of listeners, and 3 different pairs). The largest order effect occurred in the responses for American English listeners responding to [l]/[n] pairs. For this group the average subjective difference was rated as 3.02 when the order was [tana]/[tala] while the average rating was 3.34 when the order was [tala]/[tana]. Interestingly, both Nanjing groups showed the same order effect. For young Nanjing listeners the [n]/[l]

rating was 2.13, while the [l]/[n] rating was 2.29. For the older Nanjing listeners the [n]/[l] rating was 1.41, while the [l]/[n] rating was 1.53. Evidently there was something about the [l]/[n] order that resulted in the stimuli sounding a little more distinct from each other, but we don't have an explanation for what might have caused this asymmetry. Two other reliable order effects were found but they are contradictory. American English listeners rated [r]/[n] pairs as slightly more different than [n]/[r] pairs, while younger Nanjing speakers had the opposite asymmetry. Because these effects are relatively small in absolute terms and somewhat mysterious we will collapse the stimulus order for our main presentation of the results.

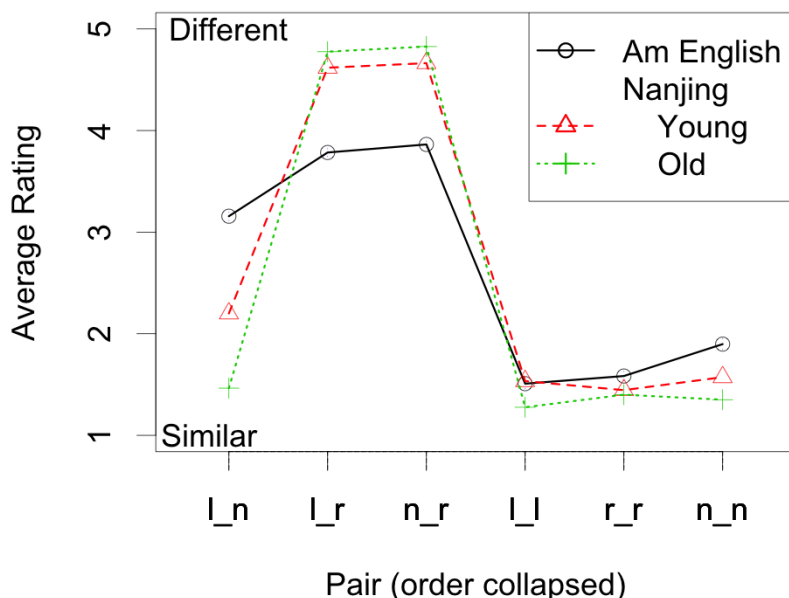


Figure 2. Average subjective distance rating for each group of listeners (speakers of American English, and younger and older speakers of the Nanjing dialect of Mandarin). Order of presentation is collapsed in this graph so that, for example, ratings of the pair [tana]/[tala] are collapsed with ratings of the pair [tala]/[tana].

The average rating values, collapsed over the order of presentation of a pair, are shown in Figure 2. For all three groups of listeners the ‘same’ pairs – comparisons of different tokens of the same intended sequence of segments – had an average similarity rating between 1 and 2 on the 5-point scale. The American English listeners appeared to use the rating scale in a different way than did the Nanjing listeners. A difference rating of “5” was rarely used by Americans for the stimuli presented in this experiment, while for Nanjing listeners “5” was the most common answer given for the [l]/[r] and [n]/[r] pairs.

All three groups of listeners rated the [l]/[n] pairs as more similar to each other than were the [l]/[r] pairs and the [n]/[r] pairs. For American English listeners the difference was rather small: 3.16 for [l]/[n] pairs versus 3.78 and 3.86 for [l]/[r] and [n]/[r] pairs. For older Nanjing listeners the difference was extreme: 1.46 for [l]/[n] pairs versus 4.78 and 4.83 for [l]/[r] and [n]/[r] pairs. The pattern for younger Nanjing listeners was similar to that for the older speakers, though with a somewhat more “different”

rating for the [l]/[n] pairs (2.2) than was found for the older speakers. There was a reliable interaction between the older and younger Nanjing listeners and pair [$X^2(2)=11.87, p<0.01$] in a linear mixed effects model with random slopes for the within subject factor ‘pairs’: rating ~ pairs * age + (pairs|subj). This model was fit to different pairs only (not pairs like [n]/[n]), and just to the Nanjing subjects. A set of post-hoc tests found that younger and older listeners’ ratings of the [l]/[n] pairs differed [$X^2(2)=9.05, p<0.01$], while the younger and older speakers ratings for the [l]/[r] and [n]/[r] pairs were not reliably different (both $p > 0.2$).

Conclusion

We replicated earlier work showing that lack of a contrast in one’s native language will have a perceptual impact (Goto, 1971; Boomershine et al., 2008; Johnson & Babel, 2010; Huang & Johnson, 2010). In addition, we showed that exposure to a non-native phonological contrast as a language learner seems to have a gradient perceptual enhancement effect on that contrast. Younger speakers in Nanjing gave ratings of the subjective similarity of [n] and [l] that were between the ratings they gave for repeated instances of the same phoneme (such as [l]/[l] pairs) and the ratings they gave for pairs that were contrastive in both Standard Mandarin and in Nanjing Mandarin (such as [n]/[r] pairs).

References

- Best, C.T.; McRoberts, G.W. and Goodell, E. (2001) Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener’s native phonological system. *J. Acoust. Soc. Am.* **109**(2), 775-794.
- Boomershine, A.; Currie Hall, K.; Hume, E. & Johnson K. (2008). The impact of allophony vs. contrast on speech perception. In P. Avery, E. Dresher & K. Rice (eds.), *Contrast in Phonology: Perception and Acquisition*. New York: Mouton de Gruyter. pp. 143-172
- Chao, Y.R. (1929) The phonology of Nanjing dialect (南京音系), *Journal of Science (科学)*, **13**(8), 1005-1036.
- Flege, J. E. (1995) Second language speech learning: Theory, findings and problems, in *Speech Perception and Linguistic Experience: Issues in Cross-language Speech Research*, edited by W. Strange. York, Timonium, MD, pp. 233-272.
- Goto, Hiromu (1971) Auditory perception by normal Japanese adults of the sounds "l" and "r." *Neuropsychologia*, **9**, 317-323.
- Huang, Tsan & Johnson, K. (2010) Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica* **67**, 243-267.
- Liao, R. & Shi, F. (1987) An experimental study on the quality of initial r in Mandarin (汉语普通话 r 声母音质的实验研究), *Studies in language and linguistics (语言研究)* **7**(2), 146-161.
- Lin, T. (1995) A study on the quality of initial “日” (日母音值考), *Journal of Yenching (燕京学报)* **1**(1), 403-421
- Johnson, K. & Babel, M. (2010) On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics* **38**, 127-136.
- Song, Yidan (2015) The phonological changes of Nanjing dialect in a century (南京方言音系百年变迁), (南京师范大学文学院学报) *Journal of the School of Chinese Language and Culture, Nanjing Normal University* **4**, 157-162.
- Werker, J. F., and Tees, R. C. (1984) Phonemic and phonetic factors in adult cross-language speech perception. *J. Acoust. Soc. Am.* **75**, 1866– 1878.