

The effect of first language (L1) dialects on the identification of Vietnamese word-final stops

Kimiko Tsukada¹, Thu T. A. Nguyễn²

¹ Department of Asian Studies, Macquarie University, Sydney, Australia

² Linguistics program, The University of Queensland, Brisbane, Australia

Kimiko.Tsukada@humn.mq.edu.au, nguyenthianhthu@email.com

Abstract

This study examined the extent to which speakers' first language (L1) dialect affects the identification of word-final stops in Vietnamese. Stops in the word-final position are unreleased in Vietnamese. Further, there is a /t/-/k/ merger in the Southern, but not the Northern dialect. We tested the hypothesis that the stop tokens produced in the Southern dialect are identified less accurately than those in the Northern dialect. The results showed that the speakers' dialect influenced the intelligibility of the final stops and the Northern dialect was more intelligible than the Southern dialect.

Index Terms: cross-language speech perception, Vietnamese, dialects

1. Introduction

In many Asian languages, word-final stops are produced without an audible release burst [1, 7]. Vietnamese is one such language [5, 6, 9, 11, 15]. In the Southern dialect (e.g., Ho Chi Min City), there is a merger of /t/ and /k/ depending on the phonetic contexts. A similar merger of /t/-/k/ and /n/-/ŋ/ in the word-final position has been reported for young Cantonese speakers [8].

Figure 1 shows formant trajectories of a Vietnamese vowel /u/ followed by /p t k/. The vowel tokens were produced by six speakers of each dialect as described in the Method section. The formant frequencies were measured at ten equidistant time points. What is most noticeable in the figure is that F2 trajectories into the final /t/ vs. /k/ are less distinct in the Southern dialect (right panel) than in the Northern dialect (left panel). It can also be seen that it is the final /t/ that affects the preceding vowel most differentially across the two dialects. In the Southern dialect, the final /t/ merges with the final /k/ and not the other way around. This type of phonemic merger may influence cross-linguistic and/or dialectal speech perception. Specifically, the /t/-/k/ merger in Vietnamese may lead to the expectation that the final stops produced by the speakers of the Southern dialect are less intelligible than those produced by the speakers of the Northern dialect.

Individuals speaking different dialects of the same language presumably have different long-term cognitive representations and respond differently to the same speech sounds. Previous research on the effect of first language (L1) dialects typically examined the perception of tones [2] and/or vowels [3, 10]. For example, even when there is a match between the speaker's and listener's L1 dialect, intelligibility of the Vietnamese tones was influenced by various factors such as the dialect of the experimenter [2]. Although the perception of consonants in second language (L2) has been investigated in various studies [4, 12-14], few studies have

examined the effect of L1 dialect on the perception of consonants (see [13], however). This may be because consonant perception is apparently categorical and may not be susceptible to the dialect effect in the same way as tones and vowels are.

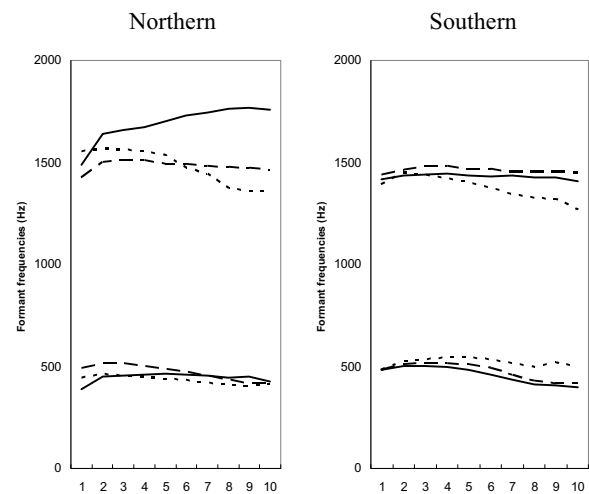


Figure 1: Formant trajectories of a Vietnamese vowel /u/ followed by /p t k/. Northern dialect (left, n=6) Southern dialect (right, n=6). The dotted lines represent vowels followed by /p/. The solid lines represent vowels followed by /t/. The dashed lines represent vowels followed by /k/.

The present study departs from previous studies by focusing on the perception of consonants. The aim of this study was to examine the effect of speakers' dialect background on the identification of final stops in Vietnamese. The key question was whether stop tokens produced in the Northern dialect are identified more accurately than those produced in the Southern dialect.

2. Method

In this experiment, the identification of Vietnamese final stops was examined by using a three-alternative forced-choice identification task as described below.

2.1. Stimuli

Monosyllabic /bVC/ words with a rising tone where V was /i a ɤ u ø/ and C was /p t k/ were produced by 12 native speakers of Vietnamese. Table 1 shows the test words used.

Table 1. *Test words used in this study.*

vowel	gloss
/i/	/bip/ ‘the sound bip’
	/bit/ ‘the first syllable of <i>bit tát</i> meaning sock’
	/bik/ ‘the second syllable of <i>viết bíc</i> meaning ballpoint pen’
/a/	/bap/ ‘non word’
	/bat/ ‘bowl’
	/bak/ ‘uncle’
/ɤ/	/bɤp/ ‘purse’
	/bɤt/ ‘to reduce the amount’
	/bɤk/ ‘a patch on the body’
/u/	/buup/ ‘non word’
	/buut/ ‘to pick (leaves)’
	/buuk/ ‘the second syllable of <i>oi búc</i> meaning sultry’
/ɐ/	/bɐp/ ‘corn’
	/bɐt/ ‘catch’
	/bɐk/ ‘north’

2.2. Speakers

Twelve native speakers of Vietnamese (6 male, 6 female) produced monosyllabic words ending with voiceless stops /p t k/. Half were speakers of the Northern dialect and the other half were speakers of the Southern dialect. Each group had the same number of male and female speakers. The recording sessions took place in the School of English, Media Studies and Art History, the University of Queensland.

2.3. Listeners

Twenty-four (7 male, 17 female) listeners with a mean age of 23.1 (range = 16 - 33) served as listeners. They were speakers of the Northern dialect and were tested in Vietnam by the second author.

2.4. Task

The listeners’ perception accuracy was tested by a three-alternative forced-choice identification task. The listeners were asked to identify the last sound of a given word by choosing one of the three stops (p t k) written in Vietnamese orthography on an answer sheet. Two tokens were excluded as they were mispronounced. Missing responses were treated as misidentification. Each stimulus item was played twice. A total of 4272 judgments (178 judgments x 24 listeners) were made.

3. Results

3.1. Overall results

Table 2 shows a confusion matrix of overall stop place identification. There was a slight tendency for bi-directional confusion between the final /t/ and /k/. On average, the listeners correctly identified the final stops 45% of the time (1922 tokens correctly identified out of 4272 tokens). This is considerably lower than the identification accuracy by both native and non-native Thai listeners previously reported [1, 12], which suggests that unreleased stops in Vietnamese may be generally less intelligible than unreleased Thai stops. Detailed cross-language acoustic comparison is needed to

confirm or disconfirm this possibility. As described below, the accuracy of identification depended both on the dialect of the speakers, the place of articulation of the final stop and the vowel preceding the final stop.

Table 2. *Confusion matrix of overall identification.*

		perceived			
		% correct	/p/	/t/	/k/
intended	/p/		47	31	21
	/t/		21	44	34
	/k/		25	29	45

A three-way ANOVA with Speakers’ Dialect (Northern, Southern), Vowel type (i a ɤ u ɐ) and Place of articulation (p t k) as within-subjects factors was carried out. The dependent variable was the mean percentages of correct identification obtained for each listener. All the main effects and two-way and three-way interaction effects were significant [D: $F(1, 23) = 72.8, p < 0.001$, V: $F(4, 92) = 3.1, p < 0.05$, P: $F(2, 46) = 20.7, p < 0.001$, D x V: $F(4, 92) = 16.4, p < 0.001$, D x P: $F(2, 46) = 5.2, p < 0.01$, V x P: $F(8, 184) = 16.7, p < 0.001$, D x V x P: $F(8, 184) = 8.5, p < 0.001$].

3.2. Effects of dialect and place of articulation

Figure 2 shows the mean identification scores by the 24 listeners according to the dialect of the speakers. The stops in the Northern dialect were more intelligible than those in the Southern dialect for all three places.

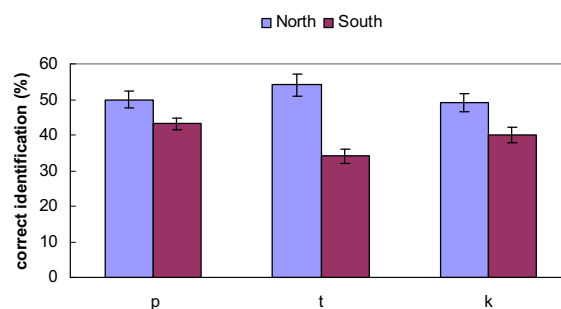


Figure 2: *Accuracy of final stop identification by 24 Vietnamese listeners. The brackets enclose ± one standard error.*

A two-way ANOVA with Speakers’ Dialect (Northern, Southern) and Place of articulation (p t k) as within-subjects factors was carried out. The main effect of Dialect reached significance [D: $F(1, 23) = 72.8, p < 0.001$], but the main effect of Place did not. The ANOVA yielded a significant two-way interaction [$F(2, 46) = 7.8, p < 0.01$], which was explored through simple effects tests.

The simple effect of Dialect was significant for all three stops [/p/: $F(1, 23) = 8.4, p < 0.01$, /t/: $F(1, 23) = 65.0, p < 0.001$, /k/: $F(1, 23) = 20.4, p < 0.001$]. The listeners correctly identified the stop place significantly more often when it was spoken by the Northern dialect speakers than when it was spoken by the Southern dialect speakers (Figure 1).

The simple effect of Place reached significance only for the stops spoken by the Southern dialect speakers [$F(2, 46) = 4.0, p < 0.05$]. The only significant difference was between /p/ (correctly identified 43% of the time) and /t/ (correctly identified 34% of the time).

Tables 3 and 4 show the confusion matrices of stop place identification according to the dialect of the speakers. It can be seen that, in the Southern dialect, the final /t/ was incorrectly perceived as /k/ more often (44%) than it was correctly perceived as /t/ (34%). In other words, the pattern of perceptual confusion can be characterized as being one-directional from /t/ to /k/. This is consistent with the observation made earlier in the Introduction on the basis of the acoustic data in Figure 1. However, as we show below, the correct identification of final stops depended on the vowel contexts.

Table 3. *Confusion matrix of stop tokens produced by Northern dialect speakers.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	50	27	22
	/t/	20	54	25
	/k/	25	25	49

Table 4. *Confusion matrix of stop tokens produced by Southern dialect speakers.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	43	36	20
	/t/	21	34	44
	/k/	25	33	40

3.3. Final stop identification in different vowel contexts

Figure 3 shows the mean identification scores by the listeners according to the dialect of the speakers for each vowel. While stop identification was more accurate in the Northern than in the Southern dialect for all vowel contexts, the extent to which final stops were accurately identified varied according to the vowel context, as well. The final stops were more intelligible in some vowel contexts (e.g., /ɜ/) than in the other (e.g., /a/). This observation was confirmed in a Dialect (2) x Vowel (5) ANOVA, which yielded significant effects for both main factors [D: $F(1, 23) = 72.8, p < 0.001$, V: $F(4, 92) = 5.8, p < 0.001$], but not for the two-way interaction effect.

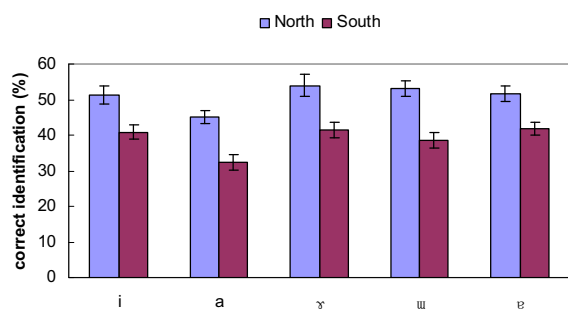


Figure 3: *Accuracy of final stop identification by 24 Vietnamese listeners. The brackets enclose ± one standard error.*

Tables 5-9 show the confusion matrices of stop place identification for the Northern dialect according to the vowel type.

Table 5. *Northern dialect: vowel /i/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	73	24	3
	/t/	44	51	6
	/k/	34	35	30

Table 6. *Northern dialect: vowel /a/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	39	26	34
	/t/	12	20	66
	/k/	15	10	73

Table 7. *Northern dialect: vowel /ɜ/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	51	15	31
	/t/	19	68	13
	/k/	28	26	42

Table 8. *Northern dialect: vowel /u/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	25	53	18
	/t/	8	81	10
	/k/	15	35	49

Table 9. *Northern dialect: vowel /ɐ/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	58	19	21
	/t/	18	45	36
	/k/	31	17	51

Tables 10-14 show the confusion matrices of stop place identification for the Southern dialect according to the vowel type. Except for when it was preceded by the /i/ and /ɜ/ vowels, the final /t/ tended to be heard predominantly as /k/ in this dialect.

Table 10. *Southern dialect: vowel /i/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	53	42	4
	/t/	33	56	10
	/k/	26	58	13

Table 11. *Southern dialect: vowel /a/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	38	41	20
	/t/	17	15	66
	/k/	31	22	44

Table 12. *Southern dialect: vowel /ɤ/.*

		perceived		
% correct		/p/	/t/	/k/
intended	/p/	49	35	15
	/t/	35	36	28
	/k/	33	26	40

Table 13. *Southern dialect: vowel /u/.*

		perceived		
% correct		/p/	/t/	/k/
	/p/	23	24	53
	/t/	9	29	62
	/k/	9	28	61

Table 14. *Southern dialect: vowel /ø/.*

		perceived		
% correct		/p/	/t/	/k/
	/p/	50	35	15
	/t/	12	33	54
	/k/	25	30	42

4. Discussion

This study examined the effect of L1 dialect on the identification of unreleased word-final stops in Vietnamese. We observed that the speakers' dialect influenced the intelligibility of the final stops in Vietnamese and the Northern dialect was more intelligible than the Southern dialect. This suggests that it may generally take longer to respond to stimuli produced by Southern dialect speakers and the extent of this delay may interact with the listeners' dialect.

An interesting question for future research is whether Southern dialect speakers of Vietnamese would be able to accurately differentiate words ending with /t/ and /k/ in their native and non-native dialects. As pointed out by one of the reviewers, this crossing of speaker and listener dialect is needed in order to discount the possibility that the superior performance of the Northern listeners with the tokens spoken in the Northern dialect is not simply attributable to the fact that they were dealing with contrasts in their own (i.e., familiar) dialect.

Vietnamese has lexical tones and, although we only included one tone (i.e., rising) in our stimuli, prosodic characteristics such as tones and voice quality may also play a role in the perception of final stops.

5. Conclusions

We have provided evidence that the /t/-/k/ merger in the word-final position in the Southern dialect of Vietnamese has perceptual relevance. To determine whether or not our results suggest an acoustical cause and generalize to other listeners, it would be necessary to examine the perception of Southern dialect speakers and to acoustically analyze the stimuli in our future work.

6. Acknowledgement

We thank three anonymous reviewers for their valuable comments on an earlier version of this paper.

7. References

- [1] Abramson, A. S. and Tingsabadh, K. (1999). Thai final stops: Cross-language perception. *Phonetica*, 56, 111-122.
- [2] Brunelle, M. and Jannedy, S. (2007). Social effects on the perception of Vietnamese tones. *Proceedings of the 16th International Congress on Phonetic Science*, Saarbrücken, Germany, 1461-1464.
- [3] Dufour, S., Nguyen, N., and Frauenfelder, U. H. (2007). The perception of phonemic contrasts in a non-native dialect. *Journal of the Acoustical Society of America*, 121, EL131-EL136.
- [4] Flege, J. E. (1989). Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, 86, 1684-1697.
- [5] Hwa-Froelich, D., Hodson, B. W., and Edward, H. T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech-Language Pathology*, 11, 264-273.
- [6] Ingram, J. C. L. and Pittam, J. (1987). Auditory and acoustic correlates of perceived accent change: Vietnamese schoolchildren acquiring Australian English. *Journal of Phonetics*, 15, 127-143.
- [7] Khouw, E. and Ciocca, V. (2006). An acoustic and perceptual study of final stops produced by profoundly hearing impaired adolescents. *Journal of Speech, Language and Hearing Research*, 49, 172-185.
- [8] Law, S.-P., Fung, R. S.-Y., and Bauer, R. (2001). Perception and production of Cantonese consonant endings. *Asia Pacific Journal of Speech, Languages and Hearing*, 6, 179-195.
- [9] Michaud, A. (2004). Final consonants and glottalization: New perspectives from Hanoi Vietnamese. *Phonetica*, 61, 119-143.
- [10] Miller, J. L. and Grosjean, F. (1997). Dialect effects in vowel perception: The role of temporal information in French. *Language and Speech*, 40, 277-288.
- [11] Tang, G. and Barlow, J. (2006). Characteristics of the sound systems of monolingual Vietnamese-speaking children with phonological impairment. *Clinical Linguistics and Phonetics*, 20, 423-445.
- [12] Tsukada, K. (2006). Cross-language perception of word-final stops in Thai and English. *Bilingualism: Language and Cognition*, 9, 309-318.
- [13] Tsukada, K., Nguyen, T. T. A., and Roengpitya, R. (2006). Cross-language perception of word-final stops by native Vietnamese listeners: Preliminary results on the role of specific, non-native phonetic experience. *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*, Auckland, New Zealand, 118-123.
- [14] Tsukada, K. and Roengpitya, R. (accepted for publication). Discrimination of English and Thai words ending with voiceless stops by native Thai listeners differing in English experience. *Journal of the International Phonetic Association*.
- [15] Vietnamese phonology (http://en.wikipedia.org/wiki/Vietnamese_phonology).