

**Univerzita Karlova**

**Filozofická fakulta**

**Fonetický ústav**

## **Teze disertační práce**

Zpracoval: Mgr. Ondřej Slówik

**Rozdíly realizace tónů hanojského a saigonského dialektu vietnamštiny mezi čteným a polospontánním mluveným projevem**

**Tone realization differences in Hanoian and Saigonese dialects between reading and semi-spontaneous speech**

vedoucí práce: doc. PhDr. Jan Volín, Ph.D.

rok odevzdání: 2018

The aspects of Vietnamese phonetics in general as well as Vietnamese tones in particular have been given scientific attention in Vietnam as well as worldwide. There are, however, certain methodological issues that need to be addressed and that triggered the origin of this thesis.

The problem of the canonical works on Vietnamese phonetics (Thompson, 1965; Đoàn, 1977; Vũ, 1982; Gordina & Bystrov 1984) is their obsolescence. Up until 1990s worldwide and well into 2000s in Vietnam, the use of computers and digital recording devices in linguistics was scarce and the data gathered by the researchers then is difficult to compare to the data we are gathering nowadays. This situation was caused simply by the state of technological development throughout the time. For this reason, researchers were forced to

rely prevalently on their hearing skills and intuition, which are tools still used in our contemporary linguistic research but no longer considered particularly reliable.

Only with the works of Nguyễn and Edmondson (1997) and later Phạm (2003) and Brunelle (2003) we can begin to talk about the rise of modern phonetic analysis using modern recording devices, recording methodology and more dependable data analysis by means of computer software. Certain objections can, nonetheless, be made towards the nature of their data. They tend to operate with small sets of speech material gathered from a limited number of speakers (usually no more than four). The recordings prevalently consist of words uttered in isolation or within artificial sentences (e.g. *Say the word "X" once*). When the authors use isolated sentences that are constructed to conform to the research purpose, they hardly feel natural in terms of collocation and context. Hence we could pose the question whether such material corresponds with real-life natural speech and whether the conclusions drawn from the gathered data represent reality accurately.

With the exception of Brunelle, who has authored several articles dealing with cross-dialect issues (comparison of Northern and Southern dialects), other prominent researchers of Vietnamese tonality have dealt predominantly with the Northern dialect and other dialects escaped their academic attention. The reason for this tendency lies in the fact that the Northern (Hanoian) is considered standard and it is also documented in greatest detail on other levels not just in terms of tonality. Moreover, two most developed departments of linguistics in Vietnam belonging to the National University are located in Hanoi and in the Ho Chi Minh City so doing research in other regions than these two is more complicated in terms of logistics, recording conditions and data storage. It is necessary to point out that there is also a significant variation in sociolinguistic consistency of the speakers. In the past 60 years, most of Vietnam has experienced turbulent population shifts. It is not difficult to

find speakers born and raised in Hanoi, whose parents were also born there or at least spent there their whole life after leaving their homeland in the nearby countryside. However, in the Ho Chi Minh City and other main cities in Southern and Central Vietnam, many people are descendants of individuals coming from the North after 1954 or 1976, which is why their accent cannot be regarded as authentically local. Consequently, selecting suitable recording subjects needs more effort and caution than in Hanoi. Researchers must be capable of accurate accent assessment otherwise the data might be compromised and might not describe the linguistic reality.

For the reasons listed above, the two dialects selected for the analysis in this thesis are the dialect of Hanoi, as it is considered standard and its speakers are considerably easy to assess, and the Ho Chi Minh City dialect where more attention had to be paid to the speaker selection in terms of family origin but it is still the dialect of the largest city in Vietnam as well as the Vietnamese economic hub. As opposed to the dialect of Hanoi that has been described in numerous academic and pedagogic works in great detail, a canonical version of the HCMC (Saigonese) dialect has not been clearly defined, which is another aspect making it necessary to approach speaker selection with utmost caution. In order to achieve highest clarity of the recorded material, it would have been optimal to compare the speech of inhabitants of two villages removed from civilization with lack of population shift. Unfortunately, it would be very bold to draw any general conclusion from the comparison of two isolated niches with limited population. Hanoi and the Ho Chi Minh City, on the other hand, represent economic and cultural centers of their regions with population amounting to millions of individuals.

Building up on the facts mentioned above, the aim of this thesis is to present the most detailed description of Hanoian and Saigonese tones in the Czech academic context.

Furthermore, the thesis is meant to attempt to improve some of the methodological strategies in researching Vietnamese tonality. The number of recorded subjects is set to 12 individuals for each dialect, which is above the average in this field of research. Furthermore, the analyzed material consists of not only isolated syllables but also reading and semi-spontaneous speech.

The beginning of the thesis, namely section 2.1. strives to introduce the topic of intonation (Cruttenden 1997; Gusehoven 2004) with focus on lexical tone in various languages across the globe (Yip 2002). The next section, 2.2., is devoted to the description of the Vietnamese language with emphasis on tonality, tonal development and dialects. More space is reserved for detailed description of the dialects spoken in Hanoi and Saigon, their segmental as well as suprasegmental structure. Chapter 3. introduces the method of speaker selection, material recording and material processing, which are key factors preparing the ground for a reliable analysis.

The empirical research part of the thesis consists of three separate tasks. The first task analyses the tones in isolation and in high/low left context in terms of pitch contours and phonation types. Its aim is to assess to what degree pitch contours of the individual tones differ across the two dialects and to what extent phonation types might be a decisive factor for tonal identification. The second part of the first task is aimed at determining the truth value of the claims made by Brunelle (2009) attesting to the presence of progressive coarticulation among the Vietnamese tones despite not playing any phonological role.

The second task encompasses a quantitative analysis of a text reading and a semi-spontaneous speech. In order to proceed with this task, it was necessary to devise a method to measure pitch values of all the analyzed syllables automatically. After applying the

method to both the read-out text and the semi-spontaneous speech, it is possible to compare the two with each other as well as across the dialects.

The third task uses a perception test to determine the ability of Hanoian and Saigonese speakers to discriminate tones across dialects and with respect to stress patterns. Respondents selected in Hanoi as well as Saigon are asked to listen to Hanoian and Saigonese syllables uttered in three levels of sentential stress and identify them. The results should provide deeper understanding of the issue of tone perception in Vietnamese.

### **Tone language**

According to Yip (2002: 17), up to 60-70 per cent of the world's languages are tonal. A language is classified as a 'tone language' if the pitch of the word can change the meaning of the word. In standard Vietnamese, for example, the syllable [ma] can be uttered with six different pitch contours to convey six different meanings.

<i>Tone</i>	<i>Meanin</i>
<i>ma</i>	Ghost
<i>mà</i>	But
<i>mã</i>	Horse
<i>mả</i>	Tomb
<i>má</i>	Cheek
<i>mạ</i>	rice

**Table 1.** *Meanings of the Vietnamese syllable [ma] in all tonal variants*

However, languages like Vietnamese or Cantonese with broad tonal registers and tones firmly embedded into syllabic cores constitute a relatively small sub-group within the large pool of tonal languages despite being considered stereotypical examples of tonal languages. Most tonal languages have smaller tonal registers and contain substantial amounts of atonal

syllables. Pike (1948) introduced two terms: ‘register tone languages’ and ‘contour tone languages’, the first type describing languages containing only level tones and the second type was a label for languages containing contour tones.

Languages such as Cantonese or Vietnamese are often described as monosyllabic, which is partially caused by orthography. Cantonese uses ideographic characters where each character normally represents one syllable and although Vietnamese uses Latin alphabet, it is actually only a transcription of an ideographic writing system based on ancient Chinese script working on very similar grounds as pin-yin. In the modern Vietnamese script (chữ quốc ngữ - national language script), every syllable is written separately and contains a tonal diacritical sign, hence the monosyllabic classification. However, there are many syllables that cannot exist on their own and they must be combined with another syllable to carry full lexical meaning, e.g. *địa* (land, earth) must be combined with other syllables like *địa lý* (geography) in order to become grammatically acceptable; the syllable *hoá* stands for the element -ize or -ization, e.g. *Tây hoá* (westernize/-ization). Neither the syllable *địa* nor *hoá* would be understood as a full lexical unit by native speakers of Vietnamese. Moreover, there are lexemes consisting of two to three syllables carrying full lexical meaning in themselves but their combination creates a new meaning, e.g. *người* (human being), *bán* (sell), *hàng* (goods) are three independent lexical units that can be combined into one - *người bán hàng* meaning ‘a vendor’. The abovementioned examples illustrate that classifying Vietnamese as a monosyllabic language does not seem to be entirely watertight. This is, nevertheless, not the concern of this dissertation.

Polysyllabic languages can either place the tone on more syllables in one word, Yip (2002: 2) mentions the example of *yùòrì* (penis) and *yúórì* (name) in Dagaare spoken in Ghana, or there can be only one tone-carrying syllable in the whole lexeme such as in the verb *ku-*

*lombéz-a* (to request) in the Chizigula language used in Tanzania. Moreover, the tone is not fixed to one syllable but can be shifted based on inflection *ku-lombež-éz-a* (to request for) or *ku-lombež-ež-án-a* (to request for each other). According to Yip (2002: 18), most polysyllabic tonal languages have more limited tonal registers and they usually contrast only two to three tones that tend to be level. As it was stated before, Vietnamese contains no atonal syllables so there are no instances similar to the Chizigula example. It might be argued that there are examples of lexemes similar to Dagaare such as *truyền thông* (media) and *truyền thống* (tradition) but that would depend on whether we pronounced Vietnamese for a truly monosyllabic language or admitted the existence of certain polysyllabic features.

### **The Vietnamese language**

Vietnamese is a tonal language belonging to the Austroasiatic language family. In the past, it used to be further classified as the Mon-Khmer group and Việt-Mường subgroup (Peiros 1998). However, Sidewell (2009) introduced the hypothesis where he claims that it is unnecessary to differentiate between Austroasiatic and Mon-Khmer rendering the labels synonymous. Khmer and Vietnamese are the only widely spoken Austroasiatic languages serving as official national languages. There are approximately 16 million Khmer speakers and around 86 million Vietnamese speakers including nearly 60 000 living in the Czech Republic.

Typologically, Čermák (2004) classified Vietnamese as a polysynthetic language because it combines full lexical items to form other lexemes e.g. *người bán hàng* (shop assistant) = *người* (person) *bán* (sell) *hàng* (goods). Full lexical items can be also used as grammar markers e.g. *mới* (new) in *Anh ấy mới về nhà*. (He has *just* gone home.). Based on isolating features such as the use of particles *đã* and *sẽ* functioning as temporal markers indicating

past and future, or the means of expressing nominal plural by adding the elements like *các* and *những* before nouns to put them in plural, Skalička (2004) classified Vietnamese as an isolating language. Typological classification of languages has been very popular in the Czech academic world. Thanks to the functionalist legacy of the Prague Linguistic Circle, it employs a very well devised methodology and it constitutes a compelling supplement or possibly even an alternative to the genealogical classification. On the other hand, the fame of typological classification of languages rarely reaches outside the scope of the functionalist tradition. Moreover, languages are organic entities and most of them contain features of multiple language types.

It is estimated (Trần Trí Dõi 2011) that about 60% of the Vietnamese vocabulary are borrowings from Chinese although the languages are genealogically rather unrelated and typologically different. Very often there are two words for one concept, one of which is considered purely Vietnamese and the other Sino-Vietnamese. Sino-Vietnamese words tend to be used in higher registers (science, academia, art, poetry, ritual and religious language) whereas the purely Vietnamese words are reserved for everyday conversation. E.g., the Vietnamese words for *wind* and *water* are *gió* and *nước*, the Sino-Vietnamese *phong* and *thuỷ*. *Phong thuỷ* is the Sino-Vietnamese transliteration of the Chinese characters 堪輿 (fēng shuǐ) and the Western world knows it as Feng-shui. Vietnamese speakers would never use the words *phong* or *thuỷ* when talking about weather. Analogically, using the syllables *gió nước* to denote Feng-shui would not be understood by native speakers.

There is a saying in Vietnamese, *phong ba bão táp không bằng ngữ pháp Việt Nam*, meaning that Vietnamese grammar is worse than a typhoon. The Vietnamese use it every time they want to emphasize how complicated and complex Vietnamese grammar is. The understanding of complexity and complicatedness can differ substantially depending on

linguistic and cultural background. It is true that Vietnamese grammar employs a lot of features that are difficult to grasp without prior understanding of the Vietnamese culture. For instance, personal pronouns are in fact nouns denoting family relationships that became grammaticalized. Vietnamese people address each other: sister, brother, father, uncle, aunt... based on age and social status despite not being related in any way. The meaning of the verb “to go” is conveyed by a set of verbs differing only by the direction of the movement. Translation of the sentence “we are going to the mountains” – *chúng ta lên núi* uses the verb *lên* meaning “to go up” whereas the verb *xuống* in *chúng ta xuống thung lũng* – “we are going to the valley” means “to go down”. There are verbs *vào* and *ra* denoting inward and outward movement as well as movement to the south and to the north. On the other hand, speakers of Indo-European languages especially of those heavily inflected will find Vietnamese grammar rather impoverished with almost non-existent morphology in terms of prefixes or suffixes, inflection, conjugation and even expressing temporality. Functions of all these missing categories are taken over by word order and various particles and elements. Memorizing paradigms and drilling their usage in speech is therefore not the main obstacle to overcome in acquisition of the Vietnamese language. However, Vietnamese is extremely rich on the lexical level with numerous cases of strict context-based partial synonymy and abundant idiomatic expressions used on daily bases across all registers.

Contrary to morphology, Vietnamese phonetics is highly complex. In terms of consonants, there are many features considered problematic to acquire by non-native speakers. There is the phonological distinction between /t/ and /th/, the velar nasal /ŋ/ in syllable-final as well as syllable-initial position, and the voiceless velar fricative /x/. Syllable-finally, we can find two noteworthy consonantal phenomena: a) voiceless plosives /p/, /t/, /k/, /c/ are unreleased lacking the explosion and therefore difficult to identify, which in some dialects leads to homophony. Moreover, these syllabic codas dramatically affect tonality as they only allow

for 2 tonal variants to occur; b) rounded vowels /u/; /o/; /ɔ/ trigger a place of articulation shift forward in velar finals /k/ and /ŋ/ so they are released as velars but finish as bilabials, hence they are called labio-velar by Phạm (2003). The vocalic system is extraordinarily rich with 9 single vowels differentiated by quality and 2 more by quantity. In addition, there are 2 semi-vowels and 3 diphthongs. The pinnacle of the Vietnamese phonetic system is the tonality. The standard dialect distinguishes 6 tones, the Saigonese distinguishes 5 and even the tonally simplest dialects of the Central Vietnam distinguish 4 tones. There has not been found any evidence for tonal sandhi or tonal reduction (Phạm, 2003), which are features typical of many tonal languages although Brunelle (2016) suggested a possible emergence of tonal sandhi in Vietnam. Moreover, Brunelle published two studies on tonal coarticulation. The first was the work on coarticulation in Northern Vietnamese (2003) and later he compared tonal coarticulation in Northern and Southern Vietnamese (2009a). His research suggests that a certain degree of bidirectional coarticulation is present in both dialects.

### **Hanoian dialect**

Yet as it was already mentioned, the reasons for higher status of the Hanoian dialect were mostly historical and political. The first king of the Nguyễn dynasty, Gia Long, moved the royal court to Huế in 1802 when he managed to unify Vietnam after a period of political turmoil into a state with borders very similar to the present situation. He was from the South so he did not have a very close relationship to Hanoi but ruling from the South would have been inconvenient as well due to the length of the country. Hence, he proclaimed Huế the capital. The last king of the Nguyễn dynasty, Bảo Đại, abdicated in 1945 (although he stayed active in the South for another decade before seeking asylum in France) and Hồ Chí Minh founded the Democratic Republic of Vietnam in the North with Hanoi as its capital city in the same year. Since Hanoi became the centre of politics as well as education of the

DRV, it was only logical for the local dialect to become the standard. In the South, the legacy of Huế still remained but after establishing the State Vietnam in 1949, Bảo Đại moved the capital to Saigon and following the Geneva Agreement in 1954 when he was forced to abdicate and flee to France, Saigon became the capital of the newly established Republic of Vietnam in 1955 with its own radio channels. The Saigonese dialect became the standard of the South. This setting, however, lasted merely two decades until the final victory of DRV and the subsequent reunification in 1976 under the new name “Socialist Republic of Vietnam”. To emphasize the victory, Saigon was renamed the Ho Chi Minh City and Hanoian Vietnamese was imposed on the Southerners as the new standard. This, of course, affected the daily linguistic behaviour of the population very marginally but the Hanoian dialect became the language of the media and show business. The strict language policy in media was loosened but the aura of the dialect spoken in Hanoi being the most prestigious still remains among the people throughout the country.

Contemporary authors (Brunelle, Phạm 2003, Kirby 2011, Kiều 2012) use the Hanoian dialect as the chief representative of the North Vietnamese dialects due to the fact that it has been described most clearly and thoroughly whereas the only comprehensive study of the Saigonese dialect was carried out by Thompson (1965) and Saigonese tones were also addressed by Vũ (1982) and Nguyễn & Edmondson (1997). There are other underlying reasons for choosing the Hanoian dialect as the research topic; mainly the easy access to suitable recording subjects as most people there speak an uncompromised variation of the dialect. Vast majority of individuals born and raised in Hanoi have parents who were also born in Hanoi or in its vicinity whereas people from the Ho Chi Minh City very often trace their immediate family roots to other places.

The Hanoian dialect is nowadays most often implied when talking about the standard Vietnamese language. The reasons for recognizing this dialect as the standard are historical

and political rather than linguistic. The main and possibly only relevant linguistic reason dwells in the fact that Hanoian Vietnamese possesses the full register of 6 tones and therefore does not manifest any signs of tonal homophony as opposed to most of the other dialects. On the downside, there is a lot of homophony in Hanoian syllable-initial consonants.

### **Saigonese dialect**

Up until 1955, when South Vietnam won independence from France, there had never been any reason for perceiving the dialect spoken in the area of the current HCM City as anything other than a regional dialect. The city of Saigon in modern proportions only came to existence in 1956 by merging the town Sài Gòn surrounding the Gia Định citadel with the Chợ Lớn market and it immediately became the capital city of the Republic of Vietnam, the capitalist and anti-communist state used by the Americans as a tool to fight Communism in South-East Asia. Although the Republic of Vietnam was defeated in 1975 and reunited with Northern Vietnam in 1976, the two decades of its existence were linguistically significant. American military presence in the Republic of Vietnam was strong, especially between 1967-1969, when it hosted more than a million American citizens. Despite the language policy focused mainly on English education of the local population, there was also need for Americans to learn Vietnamese be it for the purposes of the military, secret service, news agencies, charities or religion. In order to teach Vietnamese as a second language, textbooks had to be designed and standard Southern Vietnamese had to be defined. After the reunification in 1976, however, Saigon was renamed to Ho Chi Minh City, public offices as well as media were dominated mostly by Northerners, and North Vietnamese was imposed as the national standard with very little space for regional dialects in the public spheres. This tendency began to change after the economic reforms of Đổi Mới in 1986 and nowadays, it is very common to hear Southern Vietnamese in media, movies and public speeches.

Another reason for the relative inability to define the Saigonese standard dialect can be spotted in its variability. It is as hard to define a truly Saigonese accent as it is to find someone actually speaking it. The considerable variability of the Saigonese accent might be caused by a much larger scale of migration in the area. Whereas Hanoi has been the centre of the Vietnamese (or Việt) culture for more than a millennium, Saigon fully emerged only in 1956 as mentioned above. Moreover, there were two huge migration waves spilling over the area of modern HCM City. The first occurred after the Geneva Accords of 1954 were signed and over a million people, more than 60% of whom were Catholic (Hansen 2009, Picard 2016), feared repercussions in the North and therefore fled to the South. Moreover, Ngô Đình Diệm, the prime minister and later the president of the Republic of Vietnam, was a devout Catholic and he actively invited followers of Catholicism to the South. Although Hansen (2009) rejected the theory that Diệm invited Catholics from the North into the suburbs of Saigon in order to construct a “ring of steel”, i.e., to surround Saigon by his affiliates to stabilize the region politically, it remains a fact that majority of the Northern refugees at that time resettled in the area of current HCM City. The second immigration wave came after the fall of Saigon in 1975 when members of the North Vietnamese Army as well as followers of the Communist ideology permanently settled in the area to consolidate it after the war. The two migration waves together with the general animosity towards anything reminiscent of the Republic of Vietnam seem to constitute persuasive causes of the rather wide variation of phonetic features in the Saigonese dialect.

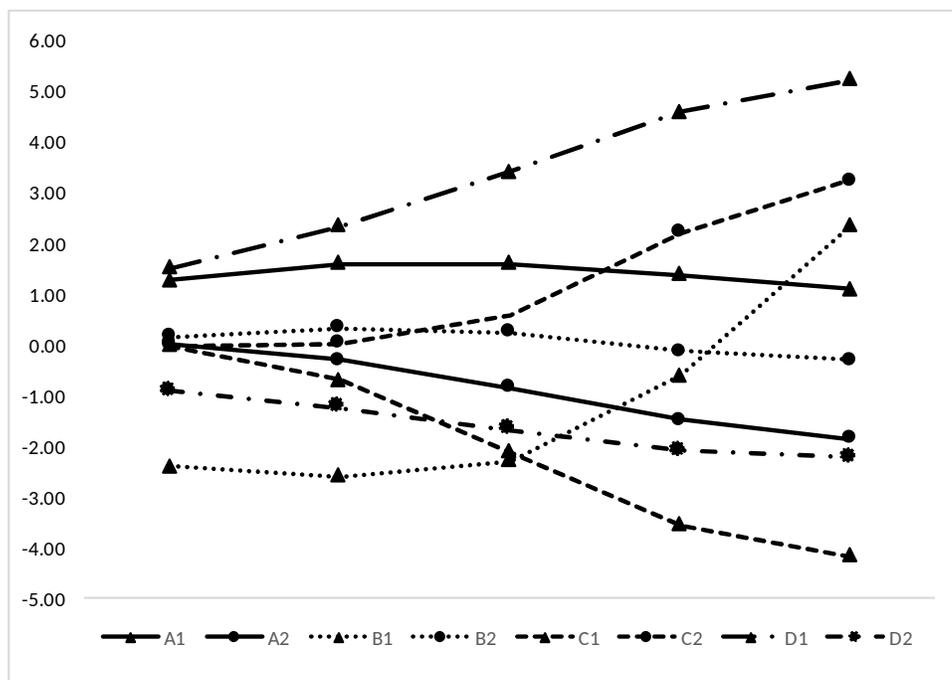
From the academic perspective, Saigonese dialect is understudied in comparison with the dialect of Hanoi. Between the mid-70s and late 90s, Saigonese dialect fell entirely out of the scope of scholarly interest. Even nowadays, the most comprehensive description of the

Saigonese dialect remains in Thompson (1965) and a few PhD theses from right after the war like Vũ (1982).

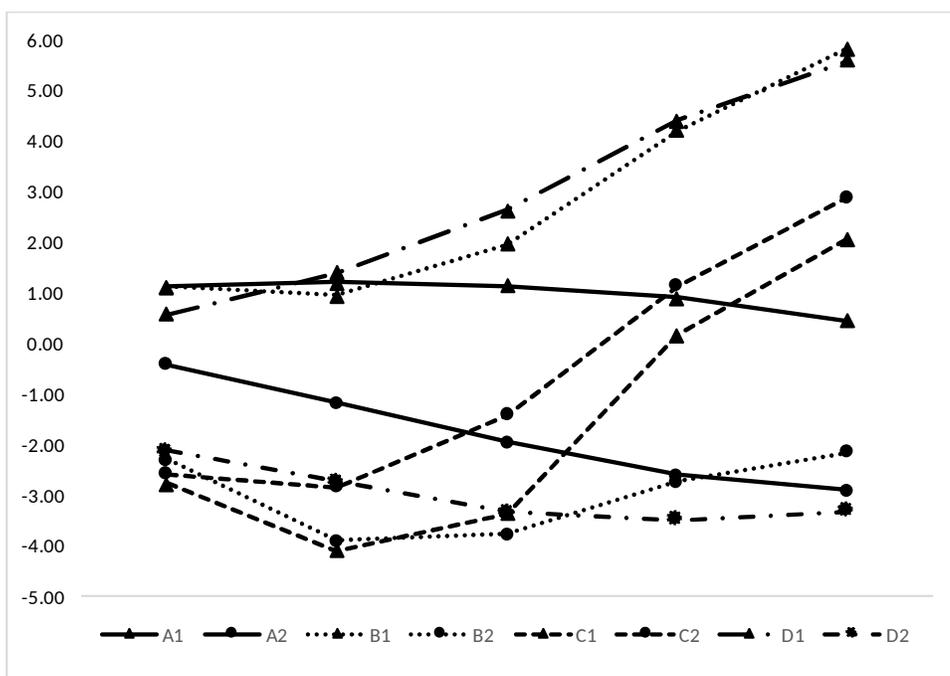
## Methodology

In order to carry out a reliable and accurate comparison of the Hanoian and Saigonese tonal systems, it was crucial to collect data that would not be compromised in terms of 1) speaker selection, 2) acoustically in the recording stage or 3) during the processing stage. Therefore, it was essential to devise dependable methodology to address all three issues and ensure data reliability. The methodology is based largely on the findings of Thomas (2002), McGuire (2010) and Niebuhr & Michaud (2015) who addressed the topics of phonetic field research – particularly speaker selection, recording conditions and linguistic use of perception tests – in substantial depth.

## Vietnamese tones in preselected syllables

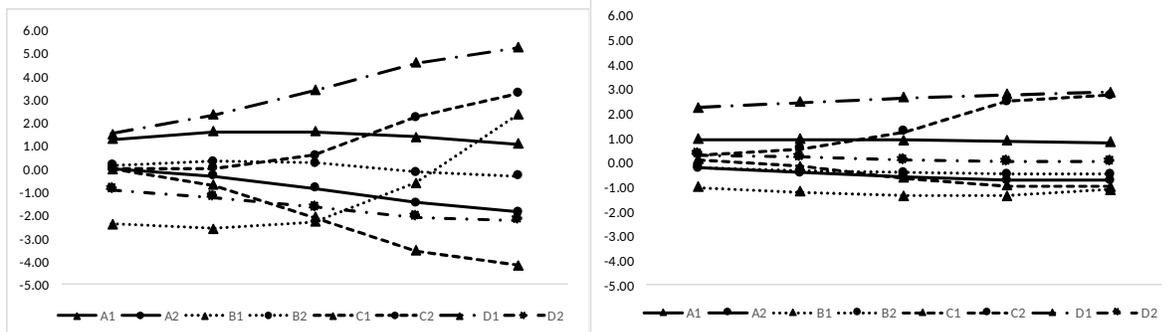


**Figure 1.** Average tone inventory of Hanoian Vietnamese. (axis y – semitones)

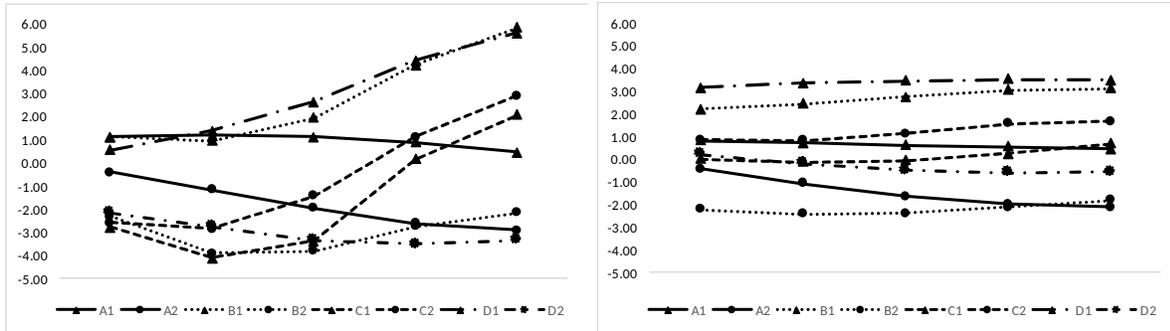


**Figure 2.** Average inventory of Saigonese Vietnamese. (axis y – semitones)

### Vietnamese tones in connected speech

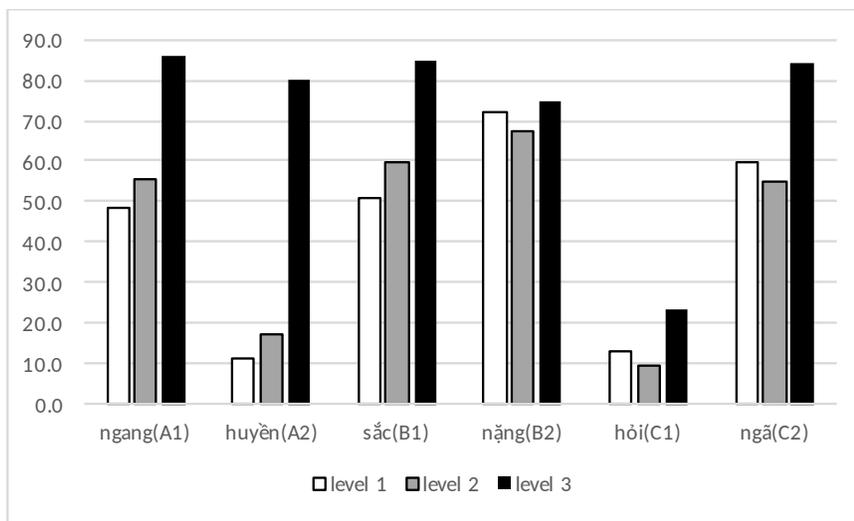


**Figure 3.** Comparison of Hanoian tone inventory representations based on “canonical realizations” (left) and on data from syllables from semi-spontaneous speech bearing stress level 1.

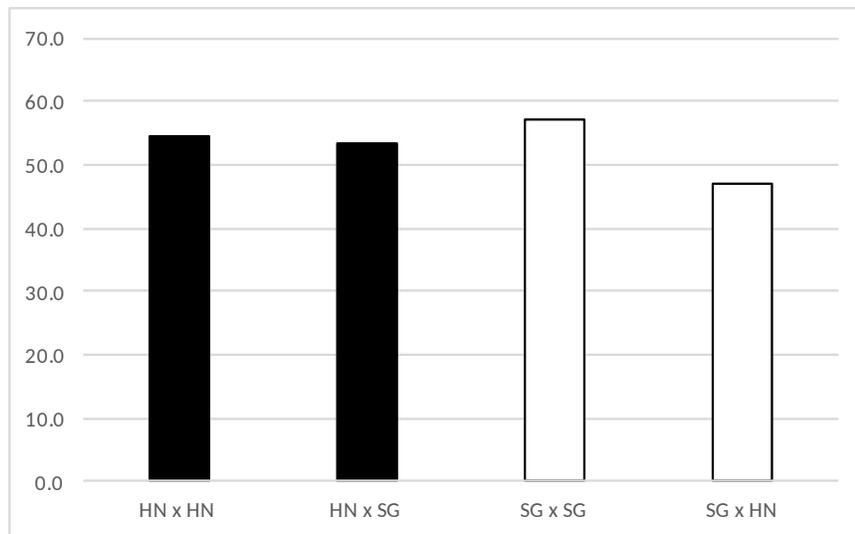


**Figure 4.** Comparison of Saigonese tone inventory representations based on “canonical realizations” (left) and on data from syllables from semi-spontaneous speech bearing stress level 1.

**Perception test**



**Figure 5.** Success rate of the individual tones in respect to the stress level (white being the weakest level and black the most prominent).



*Figure 6. Success rate of respondents within and across dialects. Hanoian speakers listening to Hanoian and Saigonese stimuli, Saigonese speakers listening to Saigonese and Hanoian stimuli.*

## **Conclusion**

It can be concluded that the two main ambitions of this dissertation have been achieved. Firstly, we managed to devise a method of  $F_0$  representation suitable for description of the Hanoian as well as the Saigonese dialect. Secondly, we created a corpus of almost 4000 tokens from 12 speakers for each dialect and added a voluminous database containing a large bulk of metadata categorized on: dialect, speaker, gender, tone, stress level, style, context, duration and  $F_0$  measurements. By doing so, we hopefully rendered the analysed data more comparable to the real-life linguistic situation.

Results gathered from the preselected syllables and syllables in isolation largely conformed to the findings of the previous studies discussed in Section 2.2. We confirmed Pham's hypothesis of 8-tone classification of Hanoian tone inventory because the contours of B1 x

D1 and B2 x D2 indeed turned out differently (even if from the systemic point of view, the idea of allotonicity should not be ruled out).

The same analysis of the Saigonese tone inventory showed a different setting. The tonal pairs in the Saigonese dialect have almost identical contour that resembles the one of Hanoian D1 x D2. We also confirmed that the contours of Saigonese tones C1 and C2 are shaped very similarly although C1 seems to be slightly lower. However, the shape seems to be superior to height as the tone C1 and C2 in HCMC are perceptually identical. Saigonese tones are also generally positioned farther from each other than the tones in Hanoi, which supports the claim by Brunelle (2009a) that discrimination of Saigonese tones depends on  $F_0$  differences more than in the case of Hanoian tones. The findings concerning Hanoian B2 and C1 have also been previously discussed but only tentatively. Although B2 is taken for a low and/or falling tone, its average  $F_0$  in this study turned out to be level and only slightly lower than high level A1. In the case of C1, it has been claimed that its contour has lost its final rise in colloquial speech but our research suggest that the rise has been lost even in careful speech. It is also noteworthy that its duration in careful annunciation is significantly shorter than of other tones with the exception of B2, D1 and D2.

The analysis of the effects of the speech style and sentential stress on tone contours indicates that both categories influence the tones in an analogical way in both dialects. With weakening stress and more spontaneous speech, the tones tend to be flatter in contour, shorter and converging to the speaker's average  $F_0$ . Figure 4.25. even suggests that the rising tone B1 in Hanoi, when occurring on syllables with weak stress or in semi-spontaneous speech, loses its properties and becomes a low and level tone below the contours of A1 and B2. In terms of duration, all tones in both dialects proportionally decrease their duration with weakening of stress and/or increasing the speech rate with the exception of Hanoian C1

that becomes equally long as A2. In this setting C1 and A2 in Hanoi are very similar in term of contour as well as duration and they might be easily confused, which partly manifested itself in Section 4.3.

The perception test was not performed utilizing as much data as tasks 4.1. and 4.2. but it nevertheless also led to certain notable observations. Firstly, respondents were having difficulties in the test administration process. It seems that abstracting tones from syllables is something that the native speakers of Vietnamese are not very comfortable with, which might have been a partial cause of their rather poor performance. What the test results reliably indicate is that the respondents were more successful discriminating tones in their own dialects. Furthermore, the Hanoian respondents scored better with stimuli from Saigonese speakers than Saigonese speakers with Hanoian stimuli. This fact can be explained by higher complexity of the tone inventory in Hanoi. Saigonese speakers using a simpler tonal inventory have greater difficulties with orientation within a more complex system. The test also indicated that Hanoian C1 was quite often misjudged as A2, which could be used as a supportive argument for possible merging of A2 and C1 in the Hanoian dialect in the future.

Finally, it is necessary to say that this dissertation has merely skimmed off the top of Vietnamese tonality as there are many more Vietnamese dialects that have not been addressed. Moreover, even the issues that have been addressed could be investigated in greater detail but that would be far beyond the scope of this dissertation. There are, however, certain issues that deserve to be addressed in the near future. Firstly, the method of  $F_0$  representation should be adjusted so that it also reflected duration of the individual tones. Secondly, subsequent research should be devoted to quantitative analysis of voice quality cues for tonal discrimination. Finally, the perception test should be replicated with more data

and under more controlled conditions to ensure that the surprisingly low success rates reflect reality and they are not merely a manifestation of a flaw in the research method.

## Bibliography

- Abramson, A. (1975) The Tones of Central Thai: Some Perceptual Experiments. In Harris and Chamberlain (eds.), *Studies in Thai Linguistics*. Bangkok: Central Institute of English Language, 1-16.
- Boersma, P. & Weenink, D. (2017). *Praat: Doing Phonetics by Computer*.
- Brunelle, M. (2003). Tone Coarticulation in Northern Vietnamese. *Proceedings of the 15th International Congress of Phonetic Sciences*. 2673-2676.
- Brunelle, M. (2009a). Northern and Southern Vietnamese Tone Coarticulation: A Comparative Case Study. *Journal of the Southeast Asian Linguistics Society*.
- Brunelle, M. (2009b). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics* 37. 79-96.
- Brunelle, M. et al. (2012). *Intonation in Northern Vietnamese*. University of Ottawa Press.
- Brunelle, M. & Phuong Hạ, Kiều & Grice, M. (2016). Inconspicuous coarticulation: A complex path to sound change in the tone system of Hanoi Vietnamese. *Journal of Phonetics*. 59.
- Cao Xuân Hạo. (2010). *Tiếng Việt, mấy vấn đề ngữ âm, ngữ pháp, ngữ nghĩa* (Vietnamese, Selected Problems from Phonetics, Grammar and Semantics). Hà Nội. NXB Giáo dục.
- Chao, Y. R. (1930). A system of tone letters. *Le maître phonétique* 45, 24-27.
- Chao, Y.R. (1968). *A grammar of spoken Chinese*. Berkeley.
- Chaudhary, C. C. (1983). Word stress in Vietnamese: A preliminary investigation. *Indian Linguistics* 44, 1-10.
- Chen, G-T. (1974). The pitch range of English and Chinese speakers, *Journal of Chinese Linguistics* 2, 159–171.
- Clumek, H. (1980). The Acquisition of Tone. In Yeni-komshian, Kavanagh and Ferguson (eds.). *Child Phonology, Vol. I, Production*. New York: Academic Press. 257-275.
- Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.
- Crystal, D. (1969). *Prosodic Systems and Intonation in English*. Cambridge: CUP.
- Cunningham, U. (2009). Phonetic Correlates of Unintelligibility in Vietnamese-accented English. *Proceedings, Fonetik 2009*. University of Stockholm.
- Cutler, A. and Chen, H. C. (1997). Lexical Tone in Cantonese: Spoken-word Processing. *Perception and Psychophysics*. Dallas.
- Čermák, F. (2004). *Jazyk a jazykověda*. Praha: Karolinum.
- Dediu, D. & Ladd, R.D. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, *ASPM* and *Microcephalin*. *Proceedings of the National Academy of Sciences Jun 2007*, 104 (26)
- Deutsch, D., Dooley, K., Henthorn, T., and Head, B. (2009). Absolute pitch among students in an American music conservatory: association with tone language fluency, *Journal of Acoustical Society of America* 125. 2398–2403.
- Đoàn Thiện Thuật. (1977). *Ngữ âm tiếng Việt* (Vietnamese Phonetics). Hà Nội: NXB Đại học và trung học chuyên nghiệp.
- Diffloth, G. (1989). Proto-Austroasiatic Creaky Voice. *Mon-Khmer Studies* 15: 139-54.
- Everett, C. et al. (2015). Climate, Vocal Folds, and Tonal Languages: Connecting the Physiological and Geographic Dots. *Proceedings of the National Academy of Sciences*, 112. 1322-27.
- Ewan, W. G. (1976). *Laryngeal Behavior in Speech*. Ph.D. Dissertation. University of California, Berkeley.
- Ferlus, M. (2004). The origin of tones in Viet-Muong, in *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society*, S. Burusphat (ed), Arizona, pp. 297-313.
- Fu, Q. J. and Zeng, F. G. (2000). Identification of Temporal Envelope Cues in Chinese Tone Recognition. *Asia Pacific Journal of Speech, Language and Hearing* 5. 45-57.

- Gordina, M. V. & Bystrov, I. S. (1984). *Фонетический строй вьетнамского языка* (Phonetic Structure of the Vietnamese Language). Moskva: Nauka.
- Greenberg, S. and Zee, E. (1979). On the Perception of Contour Tones. *UCLA Working Papers in Phonetics* 45. 150-165.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University.
- Harris, M.S. and N. Umeda (1987). Difference Limens for Fundamental Frequency Contours in Sentences. *Journal of the Acoustical Society of America* 81. 1139-1145.
- Harrison, P. A. (2000). Acquiring the Phonology of Lexical Tone in Infancy. *Lingua* 110. 581-616.
- Haudricourt, André-Georges. (1954). De l'origine des tons en vietnamien. *Journal Asiatique* 242: 69–82.
- Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. The University of Chicago Press.
- Healy, D. (2004). *Teach Yourself Vietnamese*. London: McGraw-Hill.
- Hermes, D. J. 2006. Stylization of Pitch Contours. In W. de Gruyter (ed.) *Methods in Empirical Prosody Research*. ISBN 978-311018856-1.
- Hillenbrand, J. M. and Gayvert, R. T. (2015). Phonetics exercises using the Alvin experiment-control software. *Journal of Speech, Language, and Hearing Research*, 1-14.
- Hoàng Cao Cường. (1986). Suy Nghĩ Thêm về Thanh Điều Tiếng Việt (More Thoughts about Vietnamese Tones). *Ngôn ngữ*. 3:19–38.
- Hoàng Thị Châu. (1989). *Tiếng Việt trên các miền đất nước – Phương ngữ* (Vietnamese Across the Region – Dialectology). Hà Nội, NXB Khoa học xã hội.
- Hombert, J. M., Ohala, J. J. and Ewan, W. G. (1979). Phonetic Explanations for the Development of Tones. *Language* 55: 37-58.
- Hữu Quỳnh & Vương Lộc. (1980). *Khái quát về lịch sử tiếng Việt và ngữ âm tiếng Việt hiện đại* (Brief History of the Vietnamese Language and Modern Vietnamese Phonetics). Hà Nội: NXB Giáo Dục.
- Johns-Lewis, C. (1986). Prosodic differentiation of discourse modes, in *Intonation in Discourse*, pp. 199–219.
- Kirby, J. P. (2010). Dialect Experience in Vietnamese Tone Perception. *Journal of the Acoustical Society of America* 127(4), 3749-3757.
- Kirby, J. P. (2011). Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association* 41(3). 381-392.
- Klatt, D. (1973). Discrimination of Fundamental Frequency Contours in Synthetic Speech Duplications for Models of Pitch Perception. *Journal of the Acoustical Society of America* 53. 8-16.
- Ladd, D. R. (1997). *Intonational phonology*. Cambridge: Cambridge University Press.
- Law, S. P. (1990). *The Syntax and Phonology of Cantonese Sentence-final Particles*. PhD Thesis. Boston University.
- Li, C. and Thompson, S. (1977). The Acquisition of Tone in Mandarin-speaking Children. *Journal of Child Language* 4. 185-199.
- Maddieson, I. (1997). Phonetic universals. In *The Handbook of Phonetic Sciences*, ed. W. Hardcastle & J. Laver. Blackwell Publishers, Oxford. 619-639.
- Maddieson, I. (2013). Tone. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/13>, Accessed on 2018-05-21.)
- Maeda, S. (1975). Electromyographic Study of Intonational Attributes. *Progress Report*. Research Laboratory of Electronic, MIT. 261-269.
- Machač, P., Skarnitzl R. (2009). *Principles of Phonetic Segmentation*. Praha: Epocha.
- Maspero, Henri. (1912). *Phonétique historique de la langue annamite: les initiales*. Bulletin de l'Ecole Française d'Extrême-Orient 12(1): 1-127.
- Matisoff, J. A. (1973). Tonogenesis in Southeast Asia. Consonant Types and Tone, In: L. M. Hyman (ed.) *Southern California Occasional Papers in Linguistics*. Los Angeles: Linguistics Program University of Southern California.

- McGuire, Grant. (2010). *A brief primer on experimental designs for speech perception research*. Laboratory Report. 77.
- Michaud, A. (2004). Final Consonants and Glottalization: New Perspectives from Hanoi Vietnamese. *Phonetica* 61. 119–146.
- Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review* 63(2). 81-97.
- Ngô Thanh Nhân. (1984). *The Syllabeme and Pattern of Word Formation in Vietnamese*. Ph.D. dissertation. New York University.
- Nguyễn, V. L. and Edmondson, J. (1997). *Tones and voice quality in Modern Northern Vietnamese: Instrumental case studies*. Mon-Khmer Studies 28: 1-18.
- Nguyễn, T. & Ingram, J. (2006). Stress, tone and word prosody in Vietnamese compounds. In: Paul Warren and Catherine I. Watson, Proceedings of the 11th Australasian International Conference on Speech Science & Technology. *Eleventh Australasian International Conference on Speech Science and Technology 2006*. Auckland, NZ (193-198). 6-8
- Niebuhr, O., & Michaud, A. (2015). *Speech Data Acquisition - The Underestimated Challenge*. Kieler Arbeiten in Linguistik und Phonetik (KALIPHO), 3, 1-42.
- Nooteboom, S. (1999). The Prosody of Speech: Melody and Rhythm. *The Handbook of Phonetic Science*. Blackwell Reference Online.
- Ohala, J. J. (1972). The physiology of tone. In: L. M. Hyman (ed.), *Consonant types and tone*. South California Occasional Papers in Linguistics (Univ. of So. Calif.) 1.1-14.
- Ohala, J.J. (1978). Production of tone. In Fromkin, V.A. (ed.), *Tone: a linguistic survey* 5-39.
- O'Grady, W., Dobrovolsky, M., & Aronoff, M. (1997). *Contemporary linguistics: an introduction*. NY: St. Martin's Press.
- Palková, Z. (1994). *Fonetika a fonologie češtiny*. Praha: Karolinum.
- Peiros, Ilia. (1998). *Comparative Linguistics in Southeast Asia*. Pacific Linguistics Series C, No. 142. Canberra: Australian National University.
- Peng, S. H. (1997). Production and Perception of Taiwanese Tones in Different Tonal and Prosodic Context. *Journal of Phonetics* 25. 371-400.
- Peng, S. H. (2000). Lexical versus 'Phonological' Representations of Mandarin Sandhi Tones. *Acquisition and the Lexicon: Papers in Laboratory Phonology V*. Cambridge: CUP. 152-167.
- Peterson, G. E. & Barney, H. L. (1952): Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24: 175–184.
- Phạm, A. H. (2003). *Vietnamese Tone: A New Analysis*. New York: Routledge.
- Phạm, B. & McLeod, S. (2016). Consonants, vowels and tones across Vietnamese dialects. *International Journal of Speech-Language Pathology*. 18:2. 122-134.
- Pierrehumbert, J. (1980) *The Phonology and Phonetics of English Intonation*. PhD thesis, MIT. Distributed 1988, Indiana University Linguistics Club.
- Pike, K. (1948). *Tone languages*. Ann Arbor: The University of Michigan Press.
- Pollák, P., Volín, J. & Skarnitzl, R. (2007). HMM-based phonetic segmentation in Praat environment. In: *Proceedings of XIIth "Speech and Computer – SPECOM 2007"*, 537–541.
- Pollack, I. (1952). "The Information of Elementary Auditory Displays". *Journal of the Acoustical Society of America* 24.6. 745-749.
- Sidwell, Paul. (2009). *Classifying the Austroasiatic languages: History and the State of the Art*. LINCOM studies in Asian linguistics, 76. Munich: Lincom Europa.
- Skalička, V. (2004). *Souborné dílo, I. díl*. (Collected Works I). Praha: Karolinum.
- Skalička, V. (2004). *Souborné dílo, II. díl*. (Collected Works II). Praha: Karolinum.
- Slavická, B. (2008). *Praktická fonetika vietnamštiny*. Praha: Karolinum.
- Smith, N. V. (1968). Tone in Ewe. *MIT Research Laboratory of Electronic Quarterly Progress Report* 88. 290-304.

- Sundberg, Johan. (1973). Data on maximum speed of pitch changes. *Quarterly Progress and Status Report 14*. Stockholm: Speech Transmission Laboratory. 39–47.
- Takefuta, Y., Jancosek, E. G., and Brunt, M. (1972). A statistical analysis of melody curves in the intonation of American English, in *Proceedings of the 7th International Congress of Phonetic Sciences*, Montreal 1971, 1035–1039.
- Thomas, Erik R. (2002). Sociophonetic Applications of Speech Perception Experiments. *American Speech, Volume 77, Number 2*. Duke University Press. pp. 115-147.
- Thompson, L. C. (1965). *A Vietnamese Grammar*. Seattle: University of Washington Press.
- Trần Trí Dõi. 2011. *Giáo trình lịch sử tiếng Việt* (Textbook to the History of the Vietnamese Language). Hà Nội: NXB giáo dục Việt Nam.
- Vũ, T. P. (1982). Phonetic Properties of Vietnamese Tones across dialects. In: D. Bradley (ed.), *Papers in Southeast Asian Linguistics*. 55-75. Sydney, Australian National University.
- Xu, Y. (1994). Production and Perception of Coarticulated Tones. *Journal of Acoustical Society of America* 95.4. 2240-2253.
- Xu, Y. (1999a). F<sub>0</sub> Peak Delay: When, Where, and Why It Occurs. In Ohala, J. (ed.), *International Congress of Phonetic Sciences 1999*. San Francisco. 1881-4.
- Xu, Y. (1999b). Effects of Tone and Focus on the Formation and Alignment of F<sub>0</sub> Contours. *Journal of Phonetics* 27. 55-105.
- Yip, M. (2002). *Tone*. Cambridge: CUP.
- Zhang, J. (2000). Phonetic duration effects on contour tone distribution. In M. Hirotani, A. Coetzee, N. Hall, and J-Y. Kim (eds.), *Proceedings of the 30th annual meeting of the North East Linguistic Society (NELS 30)*. GLSA Publications, Amherst, MA.