THE LEXICAL TONES
OF VIETNAMESE
METROPOLES

ONDŘEJ **SLÓWIK**
JAN **VOLÍN**

The lexical tones
of Vietnamese metropoles

Ondřej Slówik
Jan Volín

――――――

CONTENTS

LIST OF ABBREVIATIONS

C	consonant
dB	decibel
DL	difference limen
DRV	Democratic Republic of Vietnam
F0	fundamental frequency of voice
H	high pitch, high tonal target
HCMC	Ho Chi Minh City
HN	Hanoi
Hz	Hertz
JND	just noticeable difference
L	low pitch, low tonal target
M	medium pitch, medium tonal target
ms	millisecond
PT	pitch tier (a descriptive object for F0 analysis and inspection)
SG	Saigon
ST	semitone
T	tone
TBU	tone bearing unit
TG	text grid (a tagging object for speech analysis)
V	vowel
Vc	vocalic core
VN	Vietnam

PREFACE

This monograph is dedicated to the Vietnamese lexical tone – a melodic phenomenon that is not very well understood by most speakers of European languages. The background knowledge for the chief topic is built in two domains. First, it is the general use of pitch in languages with special focus on lexical tone principles. Second, the essential segmental, syllabic and prosodic features of the Vietnamese language are introduced. The core of the monograph comprises studies of lexical tone production and perception that contrast Hanoian and Saigonese tone systems. Thousands of individual instances of tones in various speech styles are analysed to uncover commonalities and differences between the two dialect. The unprecedented extent of the material should provide a solid basis for further experimenting with Vietnamese tones, whether in the area of cognition, language structure or acoustics. The deeper purpose of this book is to enrich the current insight into the nature of the language as the most significant instrument of human interaction. We wrote this text with the belief that greater appreciation of principles of language use can gradually make our lives less erratic and our actions more gratifying.

1. RATIONALE

To most speakers of Indo-European languages, the concept of a *word* stands for a meaningful language unit that serves as a building block of utterances and, with regard to its own composition, it is a string of elementary sounds that can be somehow represented by letters. In Czech, for instance, assembling elementary sounds represented by letters *v – o – d – a* will produce the word *voda* /voda/, which means *water*. Yet, for thousands of languages outside Europe (and a few in Europe itself), putting together a string of elementary sounds is just not enough. A word in these languages must have a particular melodic specification to be complete and unambiguous. Such a melodic specification is termed the *tone*[1] and it is an essential component of the word – without it the word is unfinished or at least ill-formed. Thus, the meaning of the Vietnamese syllable *ma* without a tone is unclear and can refer to, among other things, *a ghost*, *horse*, *tomb* or *mother*. Languages whose words require a tone to be properly formed are called *tone languages* (but *cf.* Section 2.1, p. 18).

Given the geographical location of the Czech Republic, tone languages might seem quite exotic and countries in which they are spoken may appear distant and irrelevant to everyday lives of ordinary citizens. Such assumption is, in fact, rather detached from reality. Although the local population of speakers of the best documented tone languages such as Mandarin, Cantonese or Thai is relatively small, there is one ethnicity of tone language speakers in the Czech Republic that is substantially large. The official number of Vietnamese people living in the country was 59 534 individuals as of September 2017[2] (www.czso.cz), which means that they constitute the third most numerous ethnic minority after the Slovaks and the Ukrainians, and the largest non-European minority in the Czech Republic.

The Vietnamese were granted the status of a national minority by law in 2013. This status entitles them, for instance, to be represented in the Government Council for National Minorities, to request funds from the Czech Ministry of Culture for propagation of the Vietnamese culture, or to obtain free counselling services in Vietnamese for the purpose of official procedures. The children are entitled to elective courses of Vietnamese language at elementary schools. As a minority the Vietnamese are generally respected as peaceful and hardworking

---

1   The full term is the *lexical tone*. In the context of this book, we will often use just *tone* for short.
2   Unofficially the number is even higher. For comparison, only 3 000 Vietnamese are estimated to live in the Slovak Republic and about the same number in Hungary, while 14 000 allegedly live in Russia. On the other hand, it is estimated that over 2,000,000 Vietnamese live in the United States of America.

people. However, successful and mutually enriching coexistence of any ethnic minority with the national majority must address the issue of the language barrier.

Although this proposition sounds very logical, the way it is tackled in the current world is not always particularly effective. The lack of true understanding of how human languages work and how they are acquired prevents teachers and educators from taking efficient steps towards multilingual societies. Current teaching methods seem to be quite cumbersome and we argue that the main obstacle to their improvement is the superficial and sometimes even incorrect insight into the nature of the language. One of the modest objectives of this book is to contribute to the field of linguistics for the sake of greater appreciation of the essence of languages.

In the Czech academic environment only one monograph relevant to our topic has been produced to the best of our knowledge. It is Vietnamese Phonetics authored by Slavická (2008) and, commendably, it follows didactic goals. It is undeniably a valuable source and a useful tool for the process of acquisition of Vietnamese as a second language, although it does not address the topic of tonality in much detail. We argue that research in the Vietnamese language and its phonetic features in particular is crucial for the reasons outlined above, but also for better understanding of how the Vietnamese acquire the Czech language, be it as a second language learnt in later stages of their lives or as a mother tongue in the case of the rising numbers of bilingual Vietnamese children growing up in the Czech cultural environment.

It has to be to mentioned at this point, however, that the Czechoslovak interest in modern Asian studies dates back to the pre-WWII era. An important event happened in 1932 when Jaroslav Průšek was granted a research scholarship at the university in Beijing. Subsequently, Průšek became the head of the newly established department[3] of Chinese studies at Charles University in Prague in 1945. The department of Vietnamese studies was established in 1960 thanks to the political aspirations of the government officials to form closer diplomatic and economic links with the Democratic Republic of Vietnam (DRV).

Throughout the history of the two organizational units (i.e., Chinese and Vietnamese), many academic works were produced, but not many dealt with linguistic issues. This trend was most probably caused by the general motivation of the students and the scope of interest on the part of the researchers. Judging from the prevailing topics of discussion panels at Asian studies conferences of the era, majority of people were making their decision to engage in Asian studies because of their interest in Asian history, literature or politics. Understandably, these topics could be satisfactorily accessed and reasonably researched even with lower linguistic proficiency in the local language than when the language itself stood for the research object.

Intriguingly, various aspects of Vietnamese phonetics in general and of Vietnamese tones in particular have been given more attention in the international academic community than in Vietnam itself. The reason for this has to be sought in the social conditions throughout the last centuries in Vietnam. Be that as it may, one of the common features of canonical works on Vietnamese phonetics (e.g., Thompson, 1965; Đoàn, 1977; Vũ, 1982; Gordina & Bystrov,

---

3   Originally, it was only a departmental section called 'seminar' and it became a legitimate department in the full sense of the word in 1950.

1984) is their obsoleteness. Up until the 1980s worldwide, and well into the new millennium in Vietnam, the use of computers, digital analytical tools and quality portable recording devices in linguistics was limited. Only exceptionally dedicated researchers found ways to obtain technological equipment of that time. Therefore, most of the data at the disposal of the researchers in the past were not particularly telling, and from today's point of view they would be actually considered insufficient. It has to be repeated that this situation was caused simply by the state of technological development, not ignorance. Researchers were forced to rely on their introspection, proprioception, and in the rare case of true speech samples, which were not very extensive and of laboratory type, on their hearing skills. Although auditory phonetics is still highly respected, it does not allow for analysis of larger, generalizable samples in the current sense. The limitations of human hearing (even if thoroughly trained) prevent the applicability of the data in technological procedures or objective evaluations across research projects.

The emergence of modern phonetic approaches in terms of recording methodology, recording devices, and more dependable data processing by means of computer software can be linked to the works of Nguyễn and Edmondson (1997), and later Phạm (2003) and Brunelle (2003). Certain objections can, nonetheless, be raised about the nature of the studied material. These authors tend to base their work on small sets of speech recordings gathered from a limited number of speakers (often just one, usually not more than four). The recordings prevalently comprise words uttered in isolation or within carrier phrases (e.g., *Say the word "X" once*). If the authors use isolated sentences at all, they seem to make them quite distant from natural speech communication in terms of collocations and contexts. Clearly, such material is valuable as a starting point in descriptions of an under-researched language, and is undoubtedly constructed to conform to the research purpose at hand, but it has to be complemented later with speech samples that better correspond with real-life discourse if conclusions drawn from the collected data are to represent reality comprehensively. In this respect we hope to contribute with our study to the broadening of linguistic knowledge in the current sense.

Another issue to be raised at the outset is the geographical, or rather dialectal delineation. With the exception of Brunelle, who has authored several articles dealing with cross-dialect comparison (northern versus southern dialects), other prominent researchers of Vietnamese tonality have dealt predominantly with the northern variety. The reason for this tendency lies in the fact that the educated northern (Hanoian) speech is considered standard. Understandably, it is also documented in greatest detail on other levels of linguistic structure and not just in terms of tonality. Moreover, two most advanced departments of linguistics in Vietnam belonging to the National University are located in Hanoi and in Ho Chi Minh City (former Saigon) so doing research in other regions than these two is less convenient in terms of logistics, and perhaps even recording conditions and data handling.

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 relocation. On the one hand, it is not difficult to find speakers born and raised in Hanoi, whose parents were also either born there, or at least spent their whole life in the city after leaving their homes in the nearby countryside. However, in Ho Chi Minh City and other main cities of southern and central Vietnam, many people are descendants of migrants coming from

the North after 1954 or 1976, which is why their accent cannot be regarded as authentically local. Consequently, selecting suitable recording subjects outside the North requires more effort and caution. Researchers must be capable of accurate accent assessment, otherwise the data might be compromised and, therefore, might not describe what they are supposed to describe.

This study focuses on the tonal differences between the Hanoian and Saigonese[4] varieties. As stated above, the dialect of Hanoi is considered standard and its speakers are relatively easy to access, whereas more attention has to be paid to the speaker selection in Ho Chi Minh City[5], which, conversely, is the largest city of Vietnam, as well as the Vietnamese economic hub. As opposed to the dialect of Hanoi that has been described in numerous academic and pedagogic publications in great detail, a canonical version of the Saigonese dialect has not been clearly defined. Hypothetically, to achieve the highest clarity of the recorded material, it would be optimal to compare speech of inhabitants of any two places with a complete absence of migration and other external influences. Unfortunately, if such places exist at all, they must be quite small. Hence, it would be unwise to draw general conclusions from the comparison of two isolated niches with limited population. Hanoi and Ho Chi Minh City, on the other hand, represent economic and cultural centres of their respective regions with population amounting to millions of individuals.[6]

To summarize, our objectives are to present a detailed description of Hanoian and Saigonese lexical tones and to open discussion on some improvements in the methodological approaches to the research of Vietnamese tonality. The number of recorded subjects is set to 24 (i.e., 12 individuals for each of the two dialects), which is comfortably above the average in this field of research. Furthermore, the analysed material consists of not only isolated syllables but also continuous read-out texts and semi-spontaneous unprepared speech. For more detail see section 4.2.

As to the layout of the present book, Chapter 2 introduces the topic of voice pitch and intonation with special focus on lexical tone in general and across various languages of the world (Cruttenden 1997; Gusenhoven 2004; Yip 2002). Chapter 3 is devoted to the description of the Vietnamese language with emphasis on tonality, tonal development and lexical tone modelling. Naturally, appropriate amount of space is reserved for a detailed description of the dialects spoken in Hanoi and Saigon, covering their segmental as well as suprasegmental structure. Chapter 4 presents the design of the empirical study that we have carried out, namely the method of speaker selection, material recording and material processing, which are the key factors preconditioning reliable analyses.

The empirical study itself comprises three separate research tasks. The first one concerns tones in isolation and in high/low melodic context. Its aim is to assess to what degree pitch contours of the individual tones differ across the two dialects, but also to what extent phonation types might serve as a decisive factor in tonal identification. The second part of the first task is aimed at the truth value of the claims made by Brunelle (2009) pertaining to

---

4   The difference between the toponyms *Ho Chi Minh City* and *Saigon* is further explained in section 2.2.3. As this monograph is oriented towards linguistics rather than politics, the terms will be used interchangeably especially because the adjective *Saigonese* already exists and it suitably complements the adjective *Hanoian*.
5   Nowadays, more than 8,000,000 people live in Ho Chi Minh City.
6   The population of Hanoi is by less than a million of inhabitants smaller than the population of Ho Chi Minh City.