



The tone system of Kim Thượng Mường: an experimental study of fundamental frequency, duration, and phonation types

Minh-Châu Nguyễn

► To cite this version:

Minh-Châu Nguyễn. The tone system of Kim Thượng Mường: an experimental study of fundamental frequency, duration, and phonation types. Linguistics. 2016. dumas-01405496

HAL Id: dumas-01405496

<https://dumas.ccsd.cnrs.fr/dumas-01405496>

Submitted on 30 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

THE TONE SYSTEM OF KIM THƯỢNG MƯỜNG: AN
EXPERIMENTAL STUDY OF FUNDAMENTAL FREQUENCY,
DURATION, AND PHONATION TYPES

MINH-CHÂU NGUYỄN

M.A. dissertation
Department of Linguistics
University of Social Sciences and Humanities
Vietnam National University

2016

Dedicated to my family and teachers.

ABSTRACT

This is a study of the tone system of the Mường dialect of Kim Thượng (province of Phú Thọ, Vietnam). Distributional analysis – a classical procedure in linguistic fieldwork – brings out five tones on smooth syllables, and two tones on stopped syllables. These tones form the topic of an experimental phonetic study. Special attention is devoted to one of the tones of smooth syllables (labelled Tone 4), which is glottalized near the beginning of the rhyme. Audio and electroglottographic recordings allow for the measurement of fundamental frequency and for an estimation of the glottal open quotient, a parameter that offers indications on phonation type. The study addresses challenges linked to the quantitative study of glottalization, in order to propose a detailed picture of how tones pattern in terms of fundamental frequency, duration, and phonation types.

TÓM TẮT

Luận văn này là một nghiên cứu trường hợp về phương ngữ Mường Kim Thượng (Phú Thọ, Việt Nam) sử dụng phương pháp phân tích phân bố, một phương pháp cổ điển trong ngôn ngữ học điền dã, để đi đến kết luận về một hệ thống thanh điệu bao gồm năm thanh chính ở các âm tiết lỏng cùng với hai thanh phụ ở các âm tiết chặt trong phương ngữ này. Đây chính là đối tượng hướng đến của nghiên cứu ngữ âm thực nghiệm. Đặc biệt trong nghiên cứu này, sự chú ý được dành cho một trong các thanh điệu chính (được kí hiệu là ‘Thanh 4’) với sự tham gia của thanh hầu hóa ở nửa đầu âm tiết. Các dữ liệu âm thanh và sóng thanh hầu ghi lại được cho phép tính toán chính xác tần số cơ bản cũng như dự đoán về độ mở thanh hầu, một tham số đáng tin cậy cung cấp thông tin về kiểu tạo âm. Đề tài được thực hiện với mong muốn vượt qua các thách thức liên quan đến nghiên cứu định lượng về thanh hầu hóa, từ đó đưa ra một gợi ý về cách mô hình hóa các chỉ số về tần số cơ bản, thời lượng và các kiểu tạo âm.

RÉSUMÉ

La présente étude porte sur le système tonal du dialecte mường de Kim Thượng (province de Phú Thọ, Vietnam). Une analyse distributionnelle (méthode classique dans l’enquête linguistique) révèle cinq tons pour les syllabes sans occlusive finale, et deux pour les syllabes à occlusive finale. Ces tons font l’objet d’une analyse phonétique expérimentale. Une attention particulière est prêtée à l’un des tons en syllabe à finale non obstruente (le ton 4), qui présente une glottalisation en début de rime. Des enregistrements audio

et électroglottographiques permettent la mesure de la fréquence fondamentale et l'estimation du quotient ouvert glottal, un paramètre qui renseigne au sujet de la qualité de voix (type de phonation). L'étude aborde certains défis liés à la quantification de la glottalisation, afin de proposer un tableau détaillé des tons.

*The Association considers that by encouraging the use
of internationally recognised phonetic symbols for all languages
its work contributes seriously to the creation and maintenance
of cordial relations between peoples of different countries.*
— *The principles of the International Phonetic Association, 1949*

ACKNOWLEDGMENTS

I am deeply grateful for help and support from all the people who in one way or another have contributed in making this study possible. I would like to express my sincere thanks to all of them.

Foremost, sincere thanks go to my two great supervisors: Prof. Trần Trí Dõi and Dr. Alexis Michaud. I wish to thank Prof. Trần Trí Dõi for helping me find my vocation, and for help with a range of issues related to fieldwork as well as my linguistics training. I am especially grateful to Dr. Alexis Michaud, for supervising this M.A. dissertation and guiding me along the path of learning about experimental phonetics. I could never thank him enough for his tireless effort and incredible patience.

My deep thanks also go to language consultants. Nguyễn Văn Chí, Sa Thị Đinh, Đinh Văn Dẫn, Sa Mạnh Hồng and many others who participated in this study made me feel at home in Kim Thượng. I am very grateful for their support and help.

Another thanks to my teachers at University of Social Sciences and Humanities, Vietnam National University (USSH-VNU), and to international experts (in particular James Kirby, Marc Brunelle, John Phan, Michel Ferlus, Guillaume Jacques, and David Holm) for their precious contribution to my training, and for their encouragements. Many thanks to the researchers and fellow students at the International Research Institute MICA (HUST).

Like many students in the field of Southeast Asian studies, I am highly indebted to the founders and curators of the SEALANG archive, which makes readily available to anyone with an internet collection a fabulous wealth of documentation about languages of the area.

Special thanks to my family for their boundless love and support over the years.

My fieldwork was partly supported by funding from the International Research Institute MICA as part of the research project “AuCo”. The Institute MICA also lent me professional recording equipment, including an electroglottograph, for my field trips. Many thanks to the Directors of the Institute and of the Speech Communication department for their generous support and encouragement.

CONTENTS

I	MAIN TEXT OF THE THESIS	1
1	INTRODUCTION	3
1.1	Glottalization as a component of lexical tone	3
1.2	Some background information about the Mường language and its speakers	9
1.2.1	The Mường as an officially recognized ethnic group	9
1.2.2	The Mường language in its Vietic context	12
1.2.3	The dialect under study: Kim Thượng Mường	14
1.3	Tone in KTM: overview and research issue	15
1.3.1	Diachronic background	15
1.3.2	The tone categories	16
1.3.3	Research topic: the phonetic templates of tones	17
2	METHOD	19
2.1	Speech materials	19
2.1.1	Experiment 1: minimal sets of real words	19
2.1.2	Experiment 2: minimal sets of non-words	22
2.2	The speakers	28
2.3	The equipment	30
2.4	File editing and annotation	33
2.5	Electroglottographic analysis	37
2.5.1	Terminology: Some basic concepts	38
2.5.2	Step-by-step description about the process of analyzing electroglottographic signal	41
3	RESULTS	51
3.1	Qualitative results for a minimal set	51
3.1.1	Speaker M1	51
3.1.2	Speaker M5	53
3.1.3	Speaker M6	54
3.1.4	Summary, and expectations about results for larger data sets	55
3.2	Towards quantification	56
3.2.1	Experiment 1	56
3.2.2	Experiment 2	61
3.2.3	Summary	68
4	DISCUSSION	71
4.1	About the tone system of KTM	71
4.2	Vietnamese dialectology, a source of inspiration for Mường studies	73
4.3	Glottalization: the linguistic roles of a phonetic universal	76
4.3.1	Introduction	76
4.3.2	Definitions	77

4.3.3	Glottalization in KTM	80
4.4	Some notes about the social dynamics of KTM	81
5	CONCLUSION	85
II	APPENDICES	87
6	APPENDIX 1: THE VOWELS AND CONSONANTS OF KTM	89
6.1	Consonants	89
6.1.1	Initial consonants	89
6.1.2	Final consonants	89
6.2	Vowels	91
7	APPENDIX 2: WORD LIST	93
8	APPENDIX 3: RECORDED MATERIALS	113
9	APPENDIX 4: SCRIPTS	115
9.1	Script for sorting the results of PeakDet	115
9.2	Scripts for plotting the data	117
9.2.1	Tone by tone	117
9.2.2	For the entire tone system	120
10	APPENDIX 5: ADDITIONAL FIGURES	127
10.1	Experiment 1	127
10.1.1	Speaker M1	127
10.1.2	Speaker M5	128
10.1.3	Speaker M6	128
10.2	Experiment 2	129
10.2.1	Each individual tone	129
10.2.2	Full tone system	132
	BIBLIOGRAPHY	134

LIST OF FIGURES

Figure 1	Vietnamese tone A ₁ (<i>ngang</i>), as realized in isolation over the syllable /kɤ/ by a speaker from Hanoi. Year of recording: 1900. 4	
Figure 2	Vietnamese tone A ₂ (<i>huyền</i>), as realized in isolation over the syllable /kɤ/ by a speaker from Hanoi. Year of recording: 1900. 5	
Figure 3	Vietnamese tone B ₁ (<i>sắc</i>), as realized in isolation over the syllable /kɤ/ by a speaker from Hanoi. Year of recording: 1900. 5	
Figure 4	Vietnamese tone B ₂ (<i>nặng</i>), as realized in isolation over the syllable /kɤ/ by a speaker from Hanoi. Year of recording: 1900. 6	
Figure 5	Vietnamese tone C ₁ (<i>hỏi</i>), as realized in isolation over the syllable /kɤ/ by a speaker from Hanoi. Year of recording: 1900. 7	
Figure 6	Vietnamese tone C ₂ (<i>ngã</i>), as realized in isolation over the syllable /kɤ/ by a speaker from Hanoi. Year of recording: 1900. 7	
Figure 7	Illustration of mistaken values at F ₀ extraction during glottalization. Voicing threshold set at 0.7 in Praat. Token: Vietnamese tone C ₂ (<i>ngã</i>); syllable /ba/; speaker from Hanoi, 1900. 8	
Figure 8	Illustration of mistaken values at F ₀ extraction during glottalization. Voicing threshold set at 0.8 in Praat. Token: Vietnamese tone C ₂ (<i>ngã</i>); syllable /ba/; speaker from Hanoi, 1900. 9	
Figure 9	Map of Vietic languages by Michel Ferlus [39]. 10	
Figure 10	Map of Vietic dialects by Nguyễn Văn Tài [111]. 11	
Figure 11	Map of KTM location in Phú Thọ province, the dialect in red (http://bando.tnmtphutho.gov.vn/map.phtml). 15	
Figure 12	The cards used for elicitation in Experiment 1. 23	
Figure 13	The scene of recording: training and performing Experiment 1. 23	
Figure 14	Experiment 2: Cards for (i) target words on the left-hand side (yellow cards for smooth syllables and pink cards for stopped syllables) and (ii) tones on the right-hand side. 27	
Figure 15	The training and performing of Experiment 2. 28	
Figure 16	Two recorders: (i) Fostex FR-2LE with two channels and (ii) Roland recorder R-44 with four channels. 30	

Figure 17	Two microphones: (i) AKG C535EB for table mic and (ii) Sennheiser HSP with MZA 900 P4 phantom power adapter for head mic. 30	
Figure 18	The electroglottographic equipment: (i) Glottal enterprises EG2-PCX with (ii) dual-channel electrodes. 31	
Figure 19	The position of equipments on the recording (an illustration from the recording of speaker F5 in August 2015). 32	
Figure 20	Two support devices: (i) Uninterruptible Power Supply (UPS Santak TG1000) and (ii) LiOA SH 10000 voltage regulator. 32	
Figure 21	A commonplace sight in Vietnam: transportation by motorbike. 33	
Figure 22	Usual annotation: examples of different treatments for voiced and unvoiced initial consonants. 36	
Figure 23	Unusual annotation: examples of treatment for two special cases: (i) missing syllable, (ii) phonologically voiceless initial consonant realized phonetically as voiced. 36	
Figure 24	Example of EGG and DEGG signals with indication of glottis closure and opening. Reproduced with permission. 39	
Figure 25	Continuum of phonation types [43, p. 384]. 40	
Figure 26	A schematic representation of data processing with PEAKDET. 42	
Figure 27	MATLAB interface when giving three parameters: (i) treatment of multiple closing peak; (ii) setting threshold for the detection of double peaks; (iii) setting number of points for EGG smoothing. 43	
Figure 28	An example of REGIONS file (.TXT) in Notepad. Experiment 1, Speaker M1 43	
Figure 29	Three figures obtained after smoothing signal and detecting peaks for one item: (i) Result of analysis: Fundamental frequency, and O_q calculated in four ways, (ii) Derivative of EGG (after smoothing), (iii) Electroglottographic signal. 44	
Figure 30	An example of good DEGG signal with well-defined closing and open peaks. 44	
Figure 31	Continuous F_0 and O_q curves as a result of good DEGG signal with well-defined closing and open peaks. 45	
Figure 32	First case of non-continuous F_0 : data of Experiment 2, Speaker M1, Item 2641. 45	
Figure 33	DEGG of non-continuous F_0 , first case: data of Experiment 2, Speaker M1, Item 2641. 46	

- Figure 34 An example of DEGG signal which shows undetected closing peaks. 46
- Figure 35 The change of F_0 curve before (left) and after (right) threshold correction. 47
- Figure 36 O_q calculated in four different ways: (i) maxima on unsmoothed DEGG signal (in green), (ii) maxima on smoothed DEGG signal (in blue), (iii) barycentre of peak on unsmoothed DEGG signal (in red), (iv) barycentre of peak on smoothed DEGG signal (in black). 47
- Figure 37 An example of O_q curves with large differences across methods as a result of imprecise opening peaks. From data of speaker M1, experiment 2, item 1331. 48
- Figure 38 An example of imprecise opening peaks: DEGG signal of item 1331, speaker M1, experiment 2. Abscissa: in samples (1 sample = 1/44,100 second). 49
- Figure 39 An example of imprecise open peaks: O_q results of item 1331, speaker M1, experiment 2. 50
- Figure 40 F_0 and O_q measurement for a minimal set illustrating the five tones over the syllable /paj/, Speaker M1. 52
- Figure 41 F_0 and O_q measurement for a minimal set illustrating the five tones over the syllable /paj/, Speaker M5. 53
- Figure 42 F_0 and O_q measurement for a minimal set illustrating the five tones over the syllable /paj/, Speaker M6. 54
- Figure 43 The average curves of fundamental frequency (no standard deviation) for five KTM tones in the case of Speaker M1, Experiment 1. 57
- Figure 44 The curves of fundamental frequency: (i) with only standard deviation (in the left); (ii) in semitones and with standard deviation (in the right) for five KTM tones in the case of Speaker M1, Experiment 1. 57
- Figure 45 Average F_0 curves of KTM tones system. Left: speaker M5. Right: speaker M6. No standard deviation. Experiment 1. 58
- Figure 46 Averaged O_q curves: (i) without standard deviation (left) and (ii) with standard deviation (right) for five KTM tones in the case of Speaker M1, Experiment 1. 59
- Figure 47 An example of glottalized Tone 4 with the open quotient curves missing values midway in the case of Speaker M6, Experiment 1: (left) Individual items: one curve per item, (right) Average data with standard deviation. 59

- Figure 48 An example of DEGG signal, with complex-repetitive patterns that do not allow for the calculation of the glottal open quotient. Speaker M1, Tone 4. 60
- Figure 49 Curves of F_0 for Tone 1 (left) and Tone 2 (right), in a data set (speaker M1, experiment 2) where these two tones were not distinguished consistently. (Individual items: one curve per item). 63
- Figure 50 O_q curves suggesting a confusion of Tone 1 and Tone 2 by speaker M1 during recording. Left: Tone 1, right: Tone 2. Experiment 2. Individual items: one curve per item. 63
- Figure 51 The curves of F_0 and O_q for Tone 3, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right (Individual items: one curve per item). 64
- Figure 52 The curves of F_0 and O_q for Tone 4, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right (Individual items: one curve per item). 65
- Figure 53 An example of creaky Tone 4 with a curve of fluctuating F_0 and corresponding DEGG signals in the cases of Speaker M1, Experiment 2, item 2931. 65
- Figure 54 The curves of F_0 and O_q for Tone 5, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right (Individual items: one curve per item). 66
- Figure 55 The curves of F_0 and O_q for Tone 6, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right. (Individual items: one curve per item) 66
- Figure 56 The curves of F_0 and O_q for Tone 7, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right. (Individual items: one curve per item) 67
- Figure 57 The averaged curves of fundamental frequency (in the left) and open quotient (in the right) for KTM tone system in the case of Speaker M1, Experiment 1 (no standard deviation), 20 tokens. 71
- Figure 58 The averaged curves of fundamental frequency (in the left) and open quotient (in the right) for KTM tone system in the case of Speaker M1, Experiment 2 (no standard deviation), 126 tokens. 72
- Figure 59 Figure of F_0 curves of Hanoi Vietnamese tones, from [61, p. 386]. 74

Figure 60 Properties characterizing different kinds of creak. A tick means that the property characterizes this type; NO means that it does not; blank means variable or unknown. Reproduced from [59]. 79

LIST OF TABLES

Table 1	Correspondences between Vietnamese and Mường for twelve body part names, adapted from [49].	12
Table 2	A brief overview of the tone system of KTM. St.V. = Standard (Hanoi) Vietnamese.	17
Table 3	The two minimal sets of experiment 1, illustrating the five tones over the syllables / paj / and / rɔ̌ /.	20
Table 4	The list of consultants, numbered sequentially according to the date of first participation in my field-work (starting in Feb. 2014).	29
Table 5	List of recorded data used in this thesis: recordings conducted on latest field trip (Feb. 2016); speakers M1, M5 and M6; two experiments. “REPET” means this is the second time that the speaker participated in the experiment.	35
Table 6	Annotation scheme.	37
Table 7	The average values of F_0 and O_q obtained by quantitative analysis of the EGG signal.	51
Table 8	The proportion of cycles for which O_q could not be estimated on the basis of the EGG signal.	60
Table 9	Initial consonants inventory	89
Table 10	Final consonants inventory	89
Table 11	Palatal final consonants c and ɲ in inherited vocabulary and in loan words	90
Table 12	Vowel inventory	91

Part I

MAIN TEXT OF THE THESIS

This volume consists of two parts: the main text of the thesis, devoted to an experimental phonetic study of tone; and Appendices that set out other data and analyses from fieldwork, including: (i) phonemic analysis (vowels and consonants), (ii) word list, (iii) recorded materials, (iv) scripts, and (v) additional figures.

INTRODUCTION

1.1 GLOTTALIZATION AS A COMPONENT OF LEXICAL TONE: AN ENDURING CHALLENGE FOR PHONETICS AND PHONOLOGY

The realization that glottalization plays a role in Vietnamese tone dates back at least to de Rhodes's 17th century Dictionary [104]; this property, which has been lost in the Southern dialect of Vietnamese, is still present in the Northern dialect [9]. The presence of one form or other of glottalization is a common characteristic of languages of the Vietic group [39].¹

From the point of view of theoretical phonology, proposing models of tone that incorporate phonation-type characteristics remains a challenge. "...[P]honation register can usually be best viewed as a "package" comprising a variety of phonatory, pitch, and other properties, and it may sometimes be difficult to determine which of these, if any, is the most basic in a linguistic or perceptual sense" [23]. There have been proposals to integrate phonation into a system of tone features, as a 'phonation register' feature; however, a review concludes that "arguments for tone features typically suffer from difficulties which make arguments for a register feature less than fully convincing" [23].

From the point of view of experimental phonetics, glottalization also constitutes a challenge. It has several acoustic correlates, such as (i) rapid changes in fundamental frequency: F_0 drops sharply when glottalization takes place, (ii) changes in signal amplitude: creaky voice has lower amplitude than modal voice, and (iii) changes in spectral slope: glottalized voicing tends to have richer harmonics than softer phonation. As an illustration of the acoustic outlook of glottalization, let us examine the six tones of smooth syllables as exemplified with the syllable *cơ* (IPA: /kɤ/), realized in the year 1900 by a native speaker from Hanoi. The recording was done at the Paris *Exposition Universelle* by Léon Azoulay, of the Paris Society for Anthropology (*Société d'Anthropologie de Paris*).²

*An audio recording
of Vietnamese tones
from the year 1900*

Figure 1 shows tone *ngang* (etymological A1). Frequencies outside the range from 600 to 6,000 Hz are attenuated, presumably due to characteristics of the recording device, and do not show on the spectrogram at the chosen level of contrast (the dynamic range chosen for drawing the spec-

1. The Vietic languages are a branch of the Austroasiatic language family. The branch was once referred to by the terms Việt-Mường, Annamese-Mường, and Vietnamuong; the term Vietic was proposed by Hayes [52] who proposed to redefine Việt-Mường as referring to a sub-branch of Vietic containing only Vietnamese and Mường. This usage is followed here.

2. The recordings can be listened to through the interface Telemeta: see <http://archives.crem-cnrs.fr/> We are grateful to the curators of the archive for granting permission to access the master digital files for the purpose of the present research.

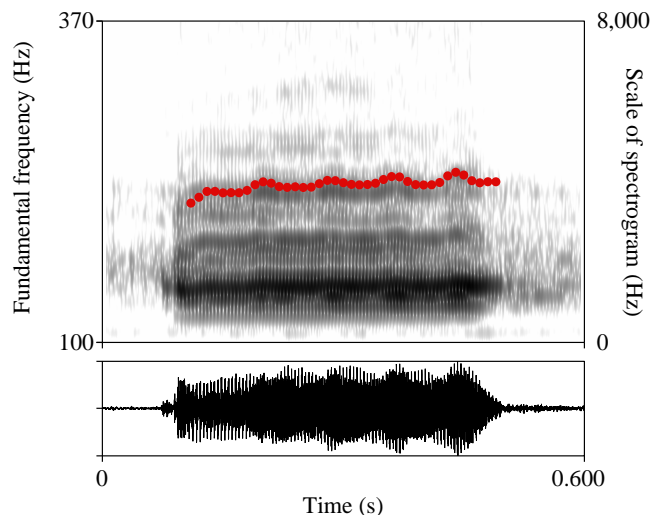


Figure 1 – Vietnamese tone A1 (*ngang*), as realized in isolation over the syllable /kx/ by a speaker from Hanoi. Year of recording: 1900.

trogram is 40 dB). Also, there is noise in the range from 1,200 to 2,500 Hz, that likewise seems attributable to the recording device. It is nonetheless possible to track fundamental frequency in these signals. Tone *ngang* (A1) looks phonetically level (relatively flat), in the upper part of the speaker's F_0 range, but not at the top of the range. The noticeable jitter and shimmer are likely to be artefacts of the recording device, although ascertaining this for sure would require more knowledge about the equipment than I have at present. Shimmer and jitter apart, this tone does not seem noticeably different than present-day realizations, more than a century later. Phonologically, this tone is not among the glottalized tones of Vietnamese; phonetic observations about this token are in keeping with this phonological property: no glottalization is discernible either on the spectrogram, on the signal, or by ear.

Tone *huyền* (etymological A2), shown in Figure 2, is falling towards the end. Again, this is not a phonologically glottalized tone, and no glottalization is discernible in the phonetic realization of this token.

The items were originally recorded in the following order: *ngang huyền nặng sắc hỏi ngã* (etymological A1, A2, B2, B1, C1, C2). For the sake of clarity of exposure I shall nonetheless follow the etymological progression: A1, A2, B1, B2, C1, C2, and thus move on presently to tone B1 (Figure 3), tone *sắc*. It has a sustained phonetic rise, straightforwardly reflecting its characterization in the phonological literature as a rising tone.

Figure 4 shows tone B2, tone *nặng*. This example introduces all at once several of the complexities associated with glottalization. The syllable shows strong changes in amplitude in the course of the vowel, pointing to glottalization in the first half of the syllable (contrary to the expectation that glottalization would be observed at the *end* of the syllable). Fundamental frequency then rises sharply, together with an increase in the signal's am-

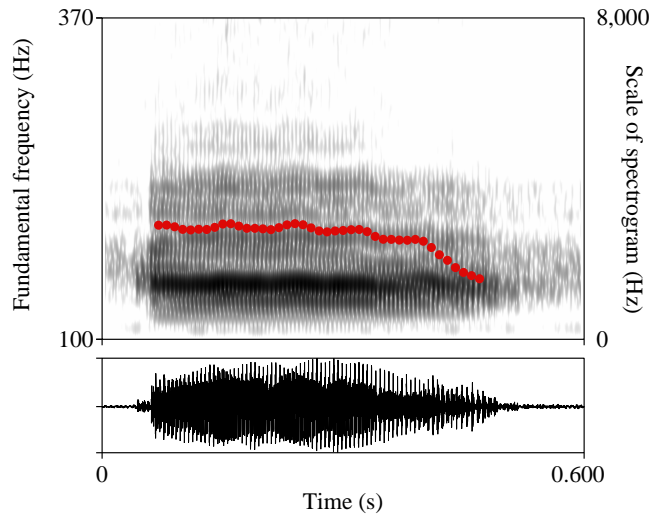


Figure 2 – Vietnamese tone A2 (*huyền*), as realized in isolation over the syllable /kʏ/ by a speaker from Hanoi. Year of recording: 1900.

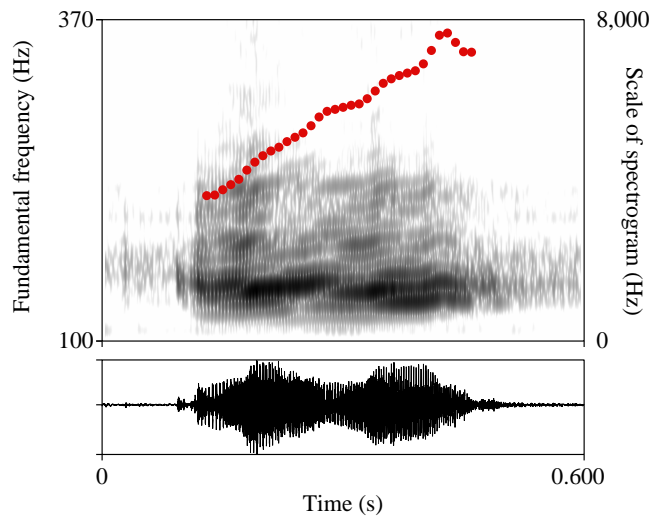


Figure 3 – Vietnamese tone B1 (*sắc*), as realized in isolation over the syllable /kʏ/ by a speaker from Hanoi. Year of recording: 1900.

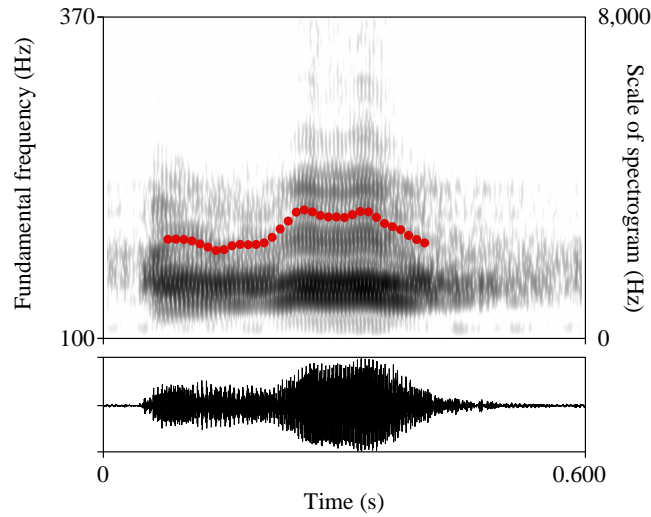


Figure 4 – Vietnamese tone B2 (*nặng*), as realized in isolation over the syllable /kx/ by a speaker from Hanoi. Year of recording: 1900.

plitude. Finally, F_0 decreases, perhaps with some final glottalization. To my ear, this token of tone B2 sounds closer to *ngã* (C2): in my dialect (standard Northern Vietnamese), a syllable with glottalization followed by a modal interval with higher pitch is categorized as *ngã* (C2), whereas tone *nặng* (B2) is associated to syllable-final glottalization. In cases where the glottalization of B2 is followed by a modal-voiced interval, its pitch must be low, otherwise the syllable will be categorized as carrying tone *ngã* (C2). If I allow myself to speculate as to why the speaker produced such a realization: it could be that this was the usual realization in his speech; but it seems more likely to me that during the rehearsal of this recording, the investigator felt that the standard realization of this tone would not be loud enough for a good recording: the syllable, cut short by glottal constriction, tends to sound as if it were strangled. It might therefore be that the person in charge of the recording instructed the consultant to produce a stronger realization, resulting in a token which is even more hyper-articulated than the other tones: a longer syllable, with more voicing after the glottalization than would usually be observed for this tone. The diversity of intonational variability of the tones serves as a word to the wise: special care is necessary at elicitation of glottalized tones.

The last two tones, C1 and C2, are shown in Figures 5 and 6 respectively. When comparing the pair formed by this speaker's B2 and C2 tones, the difference becomes aurally clear, because his tone C2 has noticeably higher overall pitch, and ends on a much higher pitch than C2.

This observation illustrates (i) the importance of the phasing of glottalization, and (ii) the phonetic relativity of the cues to tone. It calls for tools to describe these phenomena with high accuracy, to progress towards phonetic and phonological models that get as close as possible to these complex phenomena.

*Realization of tone
nặng (B2) in 1900:
was this the
canonical realization
at the time, or did an
artefact creep in?*

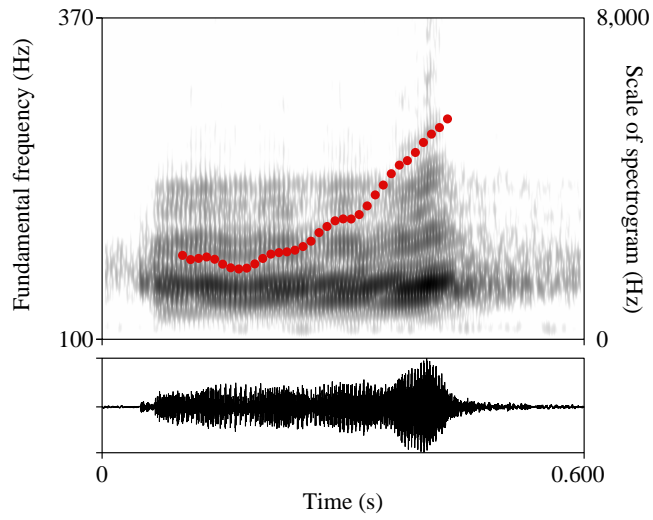


Figure 5 – Vietnamese tone C1 (*hỏi*), as realized in isolation over the syllable /kʏ/ by a speaker from Hanoi. Year of recording: 1900.

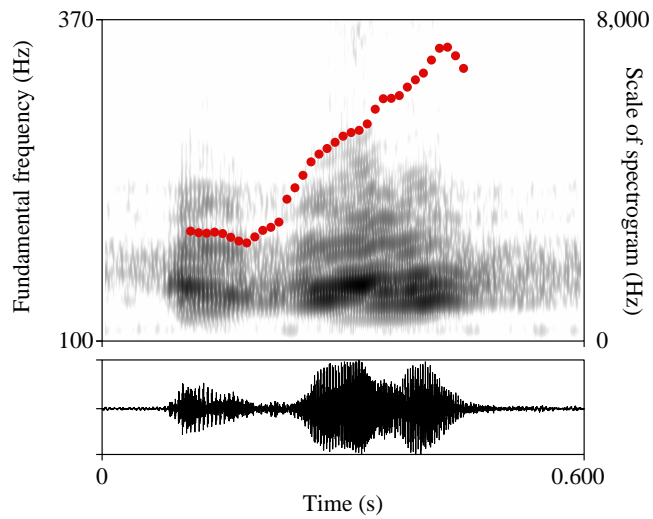


Figure 6 – Vietnamese tone C2 (*ngã*), as realized in isolation over the syllable /kʏ/ by a speaker from Hanoi. Year of recording: 1900.

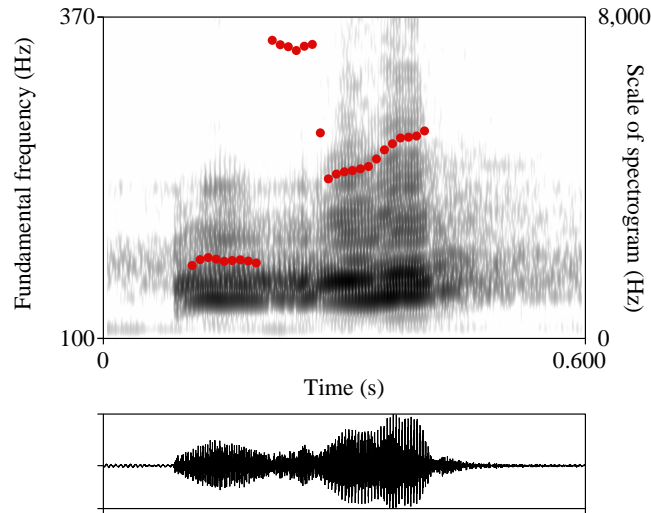


Figure 7 – Illustration of mistaken values at F_0 extraction during glottalization. Voicing threshold set at 0.7 in Praat. Token: Vietnamese tone C2 (*ngã*); syllable /*ba*/; speaker from Hanoi, 1900.

In this endeavour, an important aspect is the exploratory technique used. F_0 detection as implemented in phonetic software packages such as PRAAT can fail for glottalized portion, because the algorithms requires, if not quasi-periodicity in the signal, at least a degree of similarity between successive glottal cycles – a requirement which is often not met in audio signals corresponding to glottalized portions of speech, as exemplified by Figure 7, also taken from the 1900 recording. The token is a realization of Vietnamese tone C2 (*ngã*) over the syllable /*ba*/.

Unless specified otherwise, “Vietnamese” refers to the Northern dialect, in its “standard” Hanoian variety.

If the “Advanced pitch setting” voicing threshold in PRAAT is set at 0.7, spurious periods of high F_0 are detected (see Figure 7). If this parameter is set at 0.8, these spurious periods disappear, but no F_0 is detected on the initial portion of the syllable either: see Figure 8. Over one hundred years after this speech signal was recorded, there exists no simple way out of this technical difficulty.³

The present study aims to contribute to this line of research by: (i) exploring a language variety which (to our knowledge) had not yet been the object of experimental investigation, and (ii) testing the application of tools currently under development at the International Research Institute MICA for handling glottalization in the phonetic study of tone, aiming at statistical treatment and quantitative representations [82, 94].

3. The software package WinPitch (<http://www.winpitch.com/>) offers a range of algorithms for calculating speech fundamental frequency from audio signals. For this particular example, WinPitch’s implementation of the AMDF method (Average Magnitude Difference Function) yields good results. My point here is not that F_0 can never be computed for glottalized syllables, but that its computation is by no means straightforward, and that this topic is well worth investigating further.

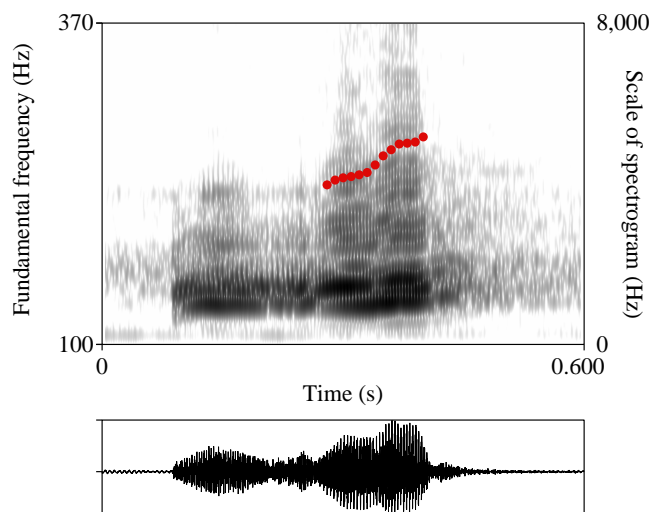


Figure 8 – Illustration of mistaken values at F_0 extraction during glottalization. Voicing threshold set at 0.8 in Praat. Token: Vietnamese tone C2 (*ngã*); syllable /*ba*/; speaker from Hanoi, 1900.

1.2 SOME BACKGROUND INFORMATION ABOUT THE MƯỜNG LANGUAGE AND ITS SPEAKERS

1.2.1 *The Mường as an officially recognized ethnic group*

As a brief introduction to the target language and its speakers, this section recapitulates census information based on administrative categories, and historical and geographical information about the Mường (Vietnamese: *người Mường*).

The Mường are currently the second largest of Vietnam's 53 officially recognized ethnic minority groups (after the Tày, speakers of a Tai-Kadai language), with an estimated population of roughly 1.3 million, based on the 2009 census and factoring in an estimation of five years of population growth [28]. Their present area of settlement spreads out over an area west, south-west, and south of the Red River [78]. They inhabit mountainous regions of northern Vietnam, with the greatest concentration in the provinces of Hòa Bình, Thanh Hóa (districts of Ngọc Lạc, Thạch Thành, Cẩm Thủy, Bá Thuộc, Như Xuân, Lang Chánh), Phú Thọ (districts of: Tân Sơn, where the fieldwork reported in this thesis was conducted; also: Thanh Sơn, Yên Lập, Tam Thanh), Sơn La (district of Phù Yên, Bắc Yên, Mộc Châu), and Nghệ An. This distribution is reflected on the map of Vietic languages by Michel Ferlus [39], reproduced here as Figure 9. In addition, Figure 10 by Nguyễn Văn Tài [111] shows an overall picture of 91 Vietic dialects in northern Vietnam and neighboring areas, with special emphasis on Mường.

Les langues du groupe viet-muong



Figure 9 – Map of Vietic languages by Michel Ferlus [39].

Mường communities are generally situated in low mountain valleys surrounded by peaks, i.e. in geographical zones contiguous with the Kinh (Vietnamese-speaking) majority, as opposed to the higher elevations typically inhabited by the Hmong or Dao. The Mường are subsistence farmers who cultivate rice (and more recently corn) as staples, alongside a number of cash crops including tea (in Phú Thọ), sugarcane (in Thanh Hóa, Phú Thọ, and Hòa Bình), and recently, acacia lumber (in Phú Thọ and Hòa Bình) [100].

The name *Mường* derives from a Tai word for ‘principality’, found in place names such as Mường Thanh, considered as the first center of Tai people when they settled in present-day Vietnam in the seventh century (in the area currently named Điện Biên Phủ). According to Trần Tử (i.e. Nguyễn Đức Tử Chi) [110], *Mường* is a term used to describe an area of Mường residents that included many villages and that was ruled by an Assembly referred to in Vietnamese as “*Nhà Lang*” or “*Thổ Lang*” (the original report is by Jeanne Cuisinier, in *Les Mường: Géographie humaine et sociologie*, 1948 [26]).⁴ No

4. *Nhà Lang* have been known as ruling class who have both rights and obligations to people in the place they control. In comparison with the Vietnamese (or Kinh) feudal landlord class, *Nhà Lang* can be assumed to have been more sensitive to the local realities. In Marxist perspective, *Nhà Lang* is praised as more “progressive”, and as representing not only itself as a class but also the whole Mường community. *Nhà Lang* had a responsibility to help Mường residents in special situations such as drought, crop failure and famine and the planning and

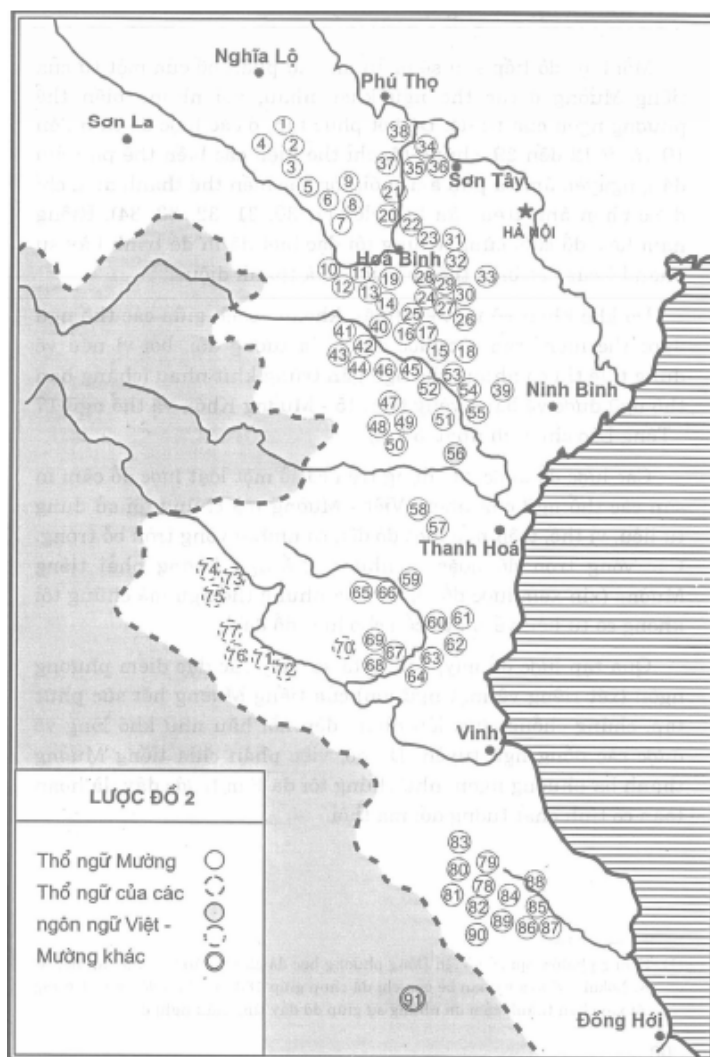


Figure 10 – Map of Vietic dialects by Nguyễn Văn Tài [111].

details are provided about the historical time window when this system was in place; it may nonetheless be relevant to the linguistic history of the area that the administration tended to be entrusted by the imperial government to local rulers who apparently belonged to the local communities and can therefore be hypothesized to be speakers of the same language as the commoners.

The people officially classified as Mường call themselves /mol/, /mwǎn/ in Hòa Bình, or /mon/, /mwǎl/ in Thanh Hóa. In Phú Thọ, especially in districts of Thanh Sơn and Tân Sơn, where the concentration of Mường people is greatest, and also in Yên Lập district and some communes of Thanh Thủy district, Mường people call themselves /mol/ or /mon/. These names are similar in terms of meaning: ‘person’; this has been analyzed as having the etymology of ‘burrower, digger’, referring to the agricultural activity

realization of major work (infrastructure work). In addition, they were in charge for the higher-level ceremonies of this region (see more in [26]).

Table 1 – Correspondences between Vietnamese and Mường for twelve body part names, adapted from [49].

numbering	English	Vietnamese (IPA)	Mường (IPA)
1	head	d̥w ^{A2} , t̥ok ^{D1}	t̥ok
2	hair	t̥ok ^{D1}	t ^h ak
3	eyes	m̥at ^{D1}	m̥at
4	ear	taj ^{A1}	t ^h ai
5	nose	muj ^{B1}	mui
6	mouth	miəŋ ^{B2}	mɛŋ
7	tooth	r̥aŋ ^{A1}	<i>not mentioned</i>
8	tongue	luəj ^{C2}	lai
9	neck	ko ^{C1}	kel, kok
10	lip	moj ^{A1}	<i>not mentioned</i>
11	chin	k̥am ^{A2}	k̥aŋ
12	arm (or hand)	t̥aj ^{A1}	t ^h ai

of horticulturists [38, p. 8]. They were borrowed into Vietnamese as “Mọi” or “Mọn”, with a derogatory meaning carrying associations of backwardness. These labels have undergone a gradual demise since the introduction of “Mường” as an official name.

1.2.2 The Mường language in its Vietic context

Mường has unmistakable similarities with Vietnamese. This was pointed out by André-Georges Haudricourt in his article “The place of Vietnamese in Austroasiatic” (English translation forthcoming) [49]. He discusses twelve words of basic vocabulary about names of body parts in Mon-Khmer languages, including: Vietnamese, Mường, Phong, Kuy, Mon, Bahnar, Mnong, and other languages. Among these, it can be seen from Table 1 that the correspondences between Vietnamese and Mường are really neat.

Mường actually provided decisive evidence for the Austroasiatic affiliation of the Vietnamese language, which had heretofore variously been considered as belonging to Tai or Chinese. There is thus no need to labor the point that studies of the Mường language have some relevance to a better understanding of the history of the Việt-Mường group: Mường dialects can yield critical evidence for the historical study of Vietnamese.

Taking a closer look at the relationship between Mường and Vietnamese – a central issue in Mường studies –, various views have been expressed. Since the Vietnamese (Kinh) and the Mường are officially two distinct ethnic groups, each tends to be considered as having its own language, hence a perception that there exists one Mường language and one Vietnamese language, each with various dialects and subdialects. The linguist’s perspective, on the

other hand, consists in taking stock of the Việt-Mường language subgroup's internal diversity, and exploring the diachronic processes of evolution that shaped this diversity, without assuming a neat two-way divide.

Within the broader Vietic subgroup, Việt-Mường in the narrow sense, also known as Northern Vietic (including Vietnamese, Mường and Nguồn) is characterized by irregular tonal reflexes in some words of basic vocabulary: the loss of proto-Austroasiatic initial voicing in Northern Vietic languages resulted in high-register reflexes, as against the expected low-register reflexes in Southern Vietic. This is interpreted by [40] as a substratum effect dating back to the time when proto-Vietic spread northwards onto an Austroasiatic substratum of languages that lacked voiced stops (a set of languages of which the Khmuic language Ksing Mul arguably constitutes a remnant). At that time, Vietnamese, Mường and Nguồn dialects were not yet distinct. The historical scenario is that their common ancestor (proto-Northern Vietic) then spread over large areas of the plains of present-day Northern Vietnam. Overall, the varieties spoken in the area of present-day Hanoi underwent more rapid evolutions, due to a variety of factors that doubtlessly included contact between various dialects and languages. Among these, one variety became dominant, and gained the status of a regional, then a national standard: the Vietnamese language. By contrast with speakers of Northern Vietic in peripheral areas, those in the area of Thăng Long (present-day Hanoi) tended to be identified by their status as townspeople, hence the name *Kinh* (the Sino-Vietnamese word for 'capital city', 京: the same that constitutes the second syllable in 'Beijing' 北京; this is also the second character found in *Tokyo*, whose name in Chinese characters is identical to one of Hanoi's former names: *Đông Kinh* 東京). This process is still active today: Frédéric Pain (p.c. 2014) observed that some proud city dwellers prefer to be identified as 'Hanoians' (*người Hà Nội*) in preference to the label 'Kinh', which is now used over the entire country as the label for the 'majority' ethnic group, and is hence less prestigious, having lost its specific association to citizenship of the capital.

Detailed ethnolinguistic study of the sense of ethnic belonging of the various communities of speakers of Northern Vietic falls outside the scope of the present study; the above remarks simply aimed to convey a feel for the gradual individuation of the Vietnamese language within Viet-Mường, a process which sheds light on the special synchronic closeness between Mường and Vietnamese.

The phonological evolution of the peripheral dialects now identified as "Mường" was less dramatic than that of Vietnamese, resulting in a more conservative phonological system, closer in some respects to proto-Northern Vietic. One of the telltale evolutions that single out Vietnamese within the Vietic subgroup is the spirantization of medial obstruents, which results in phonological correspondences such as Mường /ka/ :: Viet /ɣa/ 'chicken'. In Mường, the proto-Vietic presyllable was lost without compensation; in Vietnamese, it caused the spirantization of *ɣ [37].

A full review of linguistic studies of Mường will not be attempted here. Let us simply mention that Milton and Muriel Barker collected abundant lexical data, made available as a *Mường – Vietnamese – English Dictionary*: an extensive wordlist with Vietnamese and English glosses [5]. In addition, Miriam Barker published a bibliography of Mường and other Vietic languages in 1993, with over 40 pages of references not only in the field of linguistic studies, but also in fields such as literature, geography and history [6]. This was a key reference in gathering documentation for the present work.

Later, Nguyễn Văn Tài collected data from an impressive range of 30 dialects [111]. This is a remarkable resource to get a feel for the diversity of dialects; on the other hand, the book suffers from fairly unreliable tonal transcriptions (Ferlus p.c. 2015, confirmed by my own observations). It is a general observation that native speakers of Vietnamese can find it difficult to cast off the perceptual filter of their mother tongue and attune their ear to different tone systems.

Within the limited scope of the present study, other references about the Mường language have not yet been taken into account.

1.2.3 *The dialect under study: Kim Thượng Mường*

The variety of Mường under study here is one of the many Mường dialects found in Phú Thọ province. The exact location of fieldwork is Kim Thượng commune, Tân Sơn district,⁵ hence the label used here for this dialect: Kim Thượng Mường, abbreviated as KTM. Over 85% of the population in this district belongs to ethnic minorities. Among them, Mường make up more than 75%, and Dao account for about 6,4%, the remaining belongs to Hmông and Tày.⁶ Locally, the nation's "majority" ethnic group, the Kinh, are thus among the smallest minorities in terms of population. The Dao and Mường each form a community with festivals and habits of its own; members of other groups, such as Kinh, Tày and Hmông people, are new settlers or married into one of the two local communities.

This is a mountainous region that is difficult to access: a relatively locked area surrounded by mountains and rivers. This can be predicted to be conducive to the preservation of linguistic diversity. To the northwest is Xuân Sơn National Park; to the southwest is Đà River, which separates this area from Hoà Bình province, which contains areas of denser Mường population (see map: Figure 11). The East and the South border on other Mường dialects of Tân Sơn district. This configuration goes a long way towards explaining why KTM is still spoken today, and has not yet been entirely replaced by Vietnamese. All the local people I encountered proudly referred to themselves as /muəŋ² tǎn²/ (in Vietnamese orthography, this can be rendered

One of the (few) acronyms used in this thesis: KTM for Kim Thượng Mường, the dialect under study

5. Tân Sơn is a new district which was separated from Thanh Sơn district and established under the resolution 61/2007/NĐ-CP of the government from 09/04/2007.

6. Source: website of Phú Thọ Office of Internal Affairs, online: <http://duan600.vn/huyen-ngheo/Huyen-Tan-Son-Phu-Tho-37/>

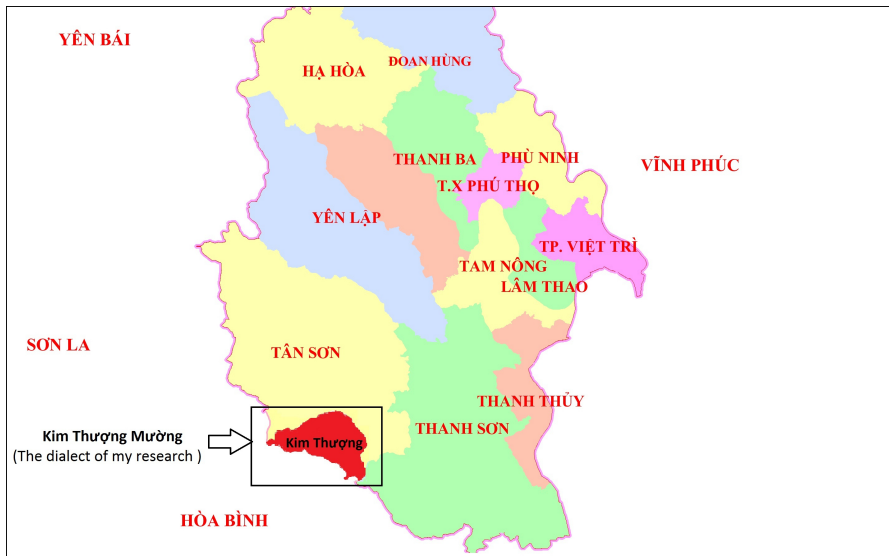


Figure 11 – Map of KTM location in Phú Thọ province, the dialect in red (<http://bando.tnmtphutho.gov.vn/map.phtml>).

as *Mường Tản*), meaning ‘the original Mường’ (from my initial field notes, it would seem that /tǎnz/ means ‘original, pristine’).

To this day, Mường is essentially an unwritten language. Some romanized writing systems were devised, and used in sizeable publications of collections of oral literature (e.g. [8]) but have not been widely adopted by the general public. Schooling is exclusively in Vietnamese, and is compulsory from age 6 to age 17.

1.3 TONE IN KTM: OVERVIEW AND RESEARCH ISSUE

1.3.1 *Diachronic background*

Tonogenesis in Vietnamese is among the success stories of 20th-century historical linguistics, and specifically, of Panchronic Phonology: an approach to diachronic phonology that searches for regularities independent of a specific language group and a specific point in time, i.e. true laws of sound change. The seminal findings of [50] and [51] are now part of the common background knowledge of specialists of Asian languages (for a summary, see e.g. [91]). A study of the historical development of tone in all the Vi-etic languages for which documentation was available at the time (1998) is proposed by [39]. This constitutes fundamental background for the present study: the system for etymological notation of tones used here (from A1 to C2 for smooth syllables, and D1 and D2 for stopped syllables) is taken up from the usual conventions set out in these diachronic studies, to which the reader is referred for further details.

1.3.2 *The tone categories*

The nature of tone itself raises issues in choosing notations for some tones of Southeast Asian languages. It seems that the tones are defined through the set of contrasts into which they enter with one another, rather than by a neat phonological property (such as the H, M and L tones of the Naxi language, for instance [92]). Transcribing this type of tone in phonetic notation remains an unsolved issue. The 1949 edition of the *Principles of the International Phonetic Association* mentioned the case of Vietnamese (referred to as ‘Annamese’): “ \sim and \vee (preceding the syllable) have been suggested for the two rising tones of Annamese, \sim implying the use of breathy voice and \vee for creaky voice” (p. 18). The intended tones are presumably B1 and C2. To my knowledge, these symbols (the macron [\sim] and the down arrowhead [\vee]) have not been taken up in later studies. An issue is that use of the macron [\sim] for one of the tones threatens to conflict with its use to indicate vowel shortness: in Vietnamese orthographic representation, \check{a} is a short vowel, contrasting with a . As orthographic representation happens to coincide with IPA transcription for these two vowels, using the macron [\sim] to indicate tone would contradict a well-established convention.

One possibility is to assign numbers to tones, without committing oneself to an analysis. This is the choice made for Mandarin Chinese, whose four tones are referred to as ‘first tone’, ‘second tone’, ‘third tone’ and ‘fourth tone’ in the linguistic literature – a practice that speaks volumes about linguists’ puzzle: the use of numbers amounts to giving up the analysis, and recognizing that there is no easy way to pinpoint the exact nature of each of the tones. Use of numbers for tones is also widespread in the field of Thai (Siamese) studies: numbering the tones of Standard Thai from one to five allows for a straightforward identification of the synchronic categories. (Interestingly, the earlier categorization of Thai tones, dating back to a time when there were only three contrastive tones, was also by means of numbers: *Mai Ek* เสียงเอก means ‘tone one’ and *Mai Tho* เสียงโท means ‘tone two’; these contrasted with a third category, left unmarked: เสียงสามัญ.) This system is adopted here: classical distributional analysis conducted during fieldwork brought out five distinctive tones for smooth syllables and two tones for stopped syllables, and these tones are numbered from 1 to 7 in Table 2.

*Assigning numbers
to tones, as a
convenient means of
identification*

Another possibility is to adopt Chao Yuen-ren’s system of five levels, to propose a stylized representation of the perceived pitch along a scale from 1 (lowest) to 5 (highest). A mid-rising tone may thus be represented as 24, 25, 34, 35, or 45. Parameters guiding the choice of one level along the five-point scale include phonetic precision and phonemic economy, which may occasionally be at odds with each other. In particular, it is recommended to use no more numbers than is phonemically necessary, i.e. use the numbers sparingly: for instance, if only three levels are necessary for the target language, these should be transcribed as 1, 3 and 5, and use of 2 and 4 would be avoided. It is for each researcher to strike a balance between economy and precision. A limitation of this system is that it does not transcribe phona-

Table 2 – A brief overview of the tone system of KTM. St.V. = Standard (Hanoi) Vietnamese.

tone	shorthand name	label	comparison with St.V.	etymology
1	Mid-level	[22]	lower than A1, <i>ngang</i>	A1
2	Falling	[52]	similar to A2, <i>huyền</i>	A2
3	Rising	[35]	similar to B1, <i>sắc</i>	B1 and B2
4	Glottalized	[2ʔ2]	similar to C1 with glottalization	C1
5	Top-high	[45]	higher but lesser sloping than B1	C2
6	Rising-checked	[24]	similar to D1	D1
7	Falling-checked	[53]	similar to D2	D2

tion types. In Hanoi Vietnamese, laryngealization plays an important role for two of the tones, namely B2 and C2; moreover, phonation-type characteristics may be found on other tones, to a greater or lesser extent (see [10]). Keeping this limitation in mind, Chao's system is nonetheless useful, and is a *de facto* standard in the description of tone in East Asian languages; for these two reasons, stylizations of the tones of KTM using Chao's system are provided in Table 2.

The description of Vietnamese tones by James Kirby [61] served as a reference in the present description. Of course the tone system of KTM is treated here as a system in its own right, not as a variant of the tone system of Vietnamese; on the other hand, Vietnamese offered a convenient point of comparison. For instance, the adoption of the label 33 in J. Kirby's description of Hanoi Vietnamese tone A1 (*ngang*) led me to adopt the label 22 for tone 1 of KTM, which is somewhat lower than Vietnamese tone A1.

1.3.3 Research topic in experimental phonetics: the phonetic templates of tones as realized in carrier sentences

The study of a phonetically complex tone system such as that of Mường presents a number of challenges. As had already been noted by Pike [101] about Mandarin Chinese, the study of tones in isolation does not tell the full story of the tone system, as tones undergo strong modification in context, due to a wealth of parameters: influences from the different vowels and consonants that compose the rhymes over which tones are realized; influences from initial consonants; coarticulation among successive tones within the utterance; and intonational modifications, reflecting phrasing and information structure.

For Vietnamese, there exists a relatively abundant literature covering the topics of (i) the templates of the lexical tones as realized in isolation or in carrier sentences [94]; (ii) tonal coarticulation [45]; and, more recently, topics of intonation [14]. Studies of Vietnamese tone set high standards for studies of tone in Mường, which should aim to vie with these achievements.

The approved order of business consists in starting out by establishing the tones' templates as realized in carrier sentences. This can then be used as a reference for the study of coarticulation and, in due course, the full range of intonational phenomena.

METHOD

The focus of this study is the phonetic templates of tones as realized in carrier sentences. Two experiments will be reported here, both targeting monosyllabic words. The purpose is to capture the main characteristics of the phonetic templates of the various lexical tones of KTM. Emphasis is placed on the full realization of tones, guiding the consultants towards a realization that is close to the *hyperarticulated* end of the continuum from hypo-articulation to hyper-articulation [80]. Use of minimal sets is ideal for the study of tones, as it factors out intrinsic properties of vowels, such as intrinsic pitch [116] and coarticulatory influences from surrounding consonants (such as co-intrinsic pitch, about which see e.g. [57]).

2.1 SPEECH MATERIALS

The target words of the experiments belong to minimal sets that contrast for tone:

- in experiment 1: minimal sets consisting of real words
- in experiment 2: minimal sets of non-words, exploring phonotactically licit combinations of initials and rhymes

The reason for combining these two experiments is that each has inherent limitations. Real words all have properties of their own: for instance, frequency of use differs across items, and this is known to have a bearing on phonetic realization. In Vietnamese, high-frequency words tend to be shorter [12]. Non-words have the advantage of perfect, abstract symmetry; they are widely used in studies of Vietnamese, for instance. On the other hand, in the absence of a writing system, they are harder to elicit. It therefore appeared interesting to try to get the best of both solutions, by setting up two distinct experiments. Proposed analyses and generalizations should ideally build on both sets of data.

Both experiments will be described in detail below. For each experiment, I will present two parts: setup and training. The *setup* paragraph explains how the experiment was designed, and how this design stems from the objectives of my study; and the *training* describes the process of explaining to the consultants the tasks in which they were invited to participate.

2.1.1 Experiment 1: minimal sets of real words

2.1.1.1 Setup

The minimal sets of tones were put together on the basis of the list of words collected in fieldwork, sorting the list by lexical tone to find sets of words that only differ in tone. As explained in section 1.3.2, KTM has five

Table 3 – The two minimal sets of experiment 1, illustrating the five tones over the syllables /paj/ and /rɔ̌/.

numbering	IPA	part of speech	Vietnamese	English
1	paj1	n	(quả) vải	lychee
	paj2	n	đập bai	barrage
	paj3	n	(cái) vại	jar
	paj4	n	trái, quả	fruit
	paj5	n	sải (tay)	armspan
2	rɔ̌1	adj	rảnh rỗi	idle
	rɔ̌2	v	no	to be sated
	rɔ̌3	v	mò (cua)	to find crab
	rɔ̌4	n	hoa chuối	banana flower
	rɔ̌5	n	(con) rùa	tortoise

tones on smooth syllables. Two minimal sets were found in the word list as of December 2015, which contained on the order of 600 words. These minimal sets are shown in Table 3,¹ and the full word list is provided in Appendix 7.

In term of positive points, all the words in these two minimal sets are common words. Using a shorter word list may be better from this point of view: all of these words are likely to be often used, as they belong to relatively basic vocabulary.

Another positive point is that none of these syllables exist in Vietnamese (segmentally). This (purely accidental) state of affairs offers a means to verify that the speakers do not enter “Vietnamese mode”: realization of /p/ as /b/ (closest Vietnamese equivalent) would reveal that the speakers have somehow shifted to Vietnamese mode.

A negative point is that different parts of speech are used. Still, all three parts of speech found in the minimal sets are content words: nouns, verbs, and adjectives. This set is likely to be relatively homogeneous, as contrasting with function words (such as pronouns or discourse particles): function words are known cross-linguistically to be realized with less articulatory energy – they are phonologically ‘weaker’, as it were. However, the difference between content and function words in Việt-Mường may be significantly lower than in most other languages. Phonologically, function words in Vietnamese and Mường have lexical tone, like content words; this is a typological difference from Mandarin Chinese dialects, for instance, where some

1. A third minimal set was initially included, and illustrated photos were prepared for the five words in this set, which appear in Figure 12. But as I went through the training stage, two of the three speakers pointed out that this third minimal set was not present in their speech. The initial fieldwork was conducted with a speaker who had a rather complex dialectal background, hence discrepancies with other speakers. Thus, I just used the two remaining sets for all speakers, so as to be able to compare across speakers.

function words are toneless (this is sometimes called ‘light tone’ in Sino-logical parlance). Phonetically, experimental research about Southern Vietnamese (based on over 60,000 syllables of conversational materials) reveals durational differences between function words and lexical words: “function words are independently shorter than lexical words”, but the difference is on an order of magnitude that is relatively limited [12]. This suggests that inclusion of function words in minimal tonal sets may be an acceptable compromise for Vietnamese (and Mường, which appears fairly similar in this respect).

A more serious shortcoming is that some of the items in the minimal sets are more monosyllabic than others from a lexical point of view. Like Vietnamese, Mường is a strictly monosyllabic language from a phonological point of view: the syllable is the domain of tone, and the language is not very far from the “morphosyllabic” type as defined by [79]: “a prototypical morphosyllabic language exhibits a fundamental unity in which syllable = morpheme = word” [102, p. 43]. On the other hand, disyllables are numerous in the lexicon. Among the words in Table 3, only one is glossed with just one monosyllable in Vietnamese, namely ‘to be sated’. For all the others, the Vietnamese gloss contains extra materials, and the Mường forms too are often accompanied by other lexical materials through collocation and ongoing disyllabification. For instance, ‘tortoise’ is *rùa* in Vietnamese, but *con rùa* is more explicit and easier to identify; likewise in Mường, /kən1 rɔ̌5/ disambiguates the root ‘tortoise’, and at elicitation, consultants who are shown an image of a tortoise may provide a disyllabic form. Part of the training therefore consisted in explaining to the speakers that the intended target item for this task is always a monosyllable.

These two minimal sets were repeated two times. That means that a total of $(5 \times 2) \times 2 = 20$ tokens were collected. They were recorded inside a carrier sentence in order to stabilize their realization. The carrier sentence is /ja2 măt6 ____ ǽŋ3/, meaning “do you know ____?” The Vietnamese translation offers a full morpheme-by-morpheme parallel to the Mường sentence: 2SG-to_know-target_item-INTERROG, “*bạn/mày biết ____ không?*”. The consultants were instructed to imagine the following context: they are teaching KTM to an intermediate learner (a non-native person) who has acquired a relatively large vocabulary. Before teaching a new word, they want to make sure that their student does not know it yet.

2.1.1.2 Training

Elicitation by means of photos appears as the best choice, as it avoids the use of another language (which could introduce a bias due to interference between languages) and also avoids the use of writing, which also introduces biases of its own. Using photos requires some preparation, however. It would be simplistic to expect that we simply show the photos and the consultant says exactly the expected word. In the two minimal sets, some words are easy to illustrate visually, in particular the nouns; but there remain words that are difficult to show by photos, such as /rɔ̌1/ ‘idle’. Hence,

I had to use several photos that do not straightforwardly point to the intended concept. In training, I explained my idea of each ambiguous image, making consultants understand and remember which image corresponds to which target word by going through practice sessions before the recording.

A drawback of this method is that the consultants focus on remembering (memory task). This may compete with the demands of good phonetic realization: it is hoped that the consultants can concentrate on clear phonetic realization of tone; if some items are more difficult to remember than others, sudden remembrance after a moment of hesitation (which can be awkward or stressful for the consultant) may be accompanied by different phonetic realization.

This is of course not to say that the task is beyond the ability of the speakers. Still, it seemed they would be more comfortable (and hence more consistent in their realization of the tones) if they were guided through the list item by item. This also allows for (approximative) control on tempo, avoiding the changes in speech rate (speeding up or slowing down) that often happen in this type of tasks (for a review, see [96]).

For this experiment, illustrated cards (photographs) as shown on Figure 12 were used to remind consultants about the target words. Each image illustrates a word, hence we have ten photos in total, arranged into two sets corresponding to the two minimal sets. Before the recording, consultants were trained. I explained photo by photo and made consultants remember which photo corresponds to which word, reminding them to say monosyllables (not disyllables), and bringing them into the imagined context that goes with the carrier sentence. Consultants were required to practise several times before recording: until they were able to go through the task with fluency. Compared to the degree of difficulty of the second experiment (about which more below), Experiment 1 did not cause great difficulty to consultants in understanding my explanation and performing as requested: the association between photos and desired target words, once explained to them, proved easy to remember. Training and recording put together took on the order of twenty minutes in total.

At recording, a fellow student who participated as an assistant showed the cards to consultants so they recognized the intended word as learnt through earlier training. Finally, they spoke out that word inside the carrier sentence. Figure 13 is a photo showing the place of recording with the board containing the photos, and the assistant holding one of the photos.

2.1.2 *Experiment 2: minimal sets of non-words, exploring phonotactically licit combinations of initials and rhymes*

2.1.2.1 *Setup*

Like any phonetic experiment, Experiment 1 (as presented in subsection 2.1.1) has some limitations: in particular, the choice of segments was dictated by lexical availability (choosing two minimal sets on the basis of a word list). In order to gain insights into intrinsic and co-intrinsic effects (coarticula-



Figure 12 – The cards used for elicitation in Experiment 1.



Figure 13 – The scene of recording: training and performing Experiment 1.

tory effects of rhymes and initials), it is useful to use different vowels and consonants. A reading experiment was therefore devised.

An obvious drawback is that the consultant's experience of reading and writing is intimately linked with their practice of the Vietnamese language. It is possible to give consultants some training in a writing system adapted from Vietnamese to Mường, and to give them the instruction to speak in Mường and not in Vietnamese. But under this setup, 'classroom' behaviour is likely to surface at various points: code-switching between the sound systems of Mường and Vietnamese, and perhaps blending both in various proportions. All consultants had received training in Vietnamese writing through primary schooling, whereas none had received training in Mường writing at school or later. It appeared interesting to attempt an experiment with reading nonetheless; comparison of the results between experiments 1 and 2 would tell to what extent differences would be found between read materials and those elicited by means of more 'ecologically valid' methods (those classically employed in fieldwork on unwritten languages).

In the first proposal of phonemes for creating target words, I selected five consonants: /p, t, k, t^h, m/. The two glottalized consonants /b/ and /d/ were excluded because they are expected to have strong coarticulatory effects on phonation type – a central topic in the study of tone. Unvoiced stops appeared as a convenient set, hence the use of /p, t, k/ (note that Mường, unlike Vietnamese, has /p/ in vocabulary of Vietic stock). The set made up of /p, t, k, t^h, m/ contained consonants from three series: unvoiced stops /p t k/, aspirated stop /t^h/, and nasal consonant /m/. An issue was that this set was not symmetrical, detracting from its usefulness in statistical analysis. It appeared better to have only two series, but at similar places of articulation, allowing for a comparison of nasals and stops. The modified set therefore consisted of six consonants: three unvoiced plosives /p t k/ corresponding to three nasals /m n ŋ/.

In term of vowels, I tried to select maximally different vowels, therefore choosing /i a u/: a high front vowel, /i/; a low vowel, /a/; and a high back (rounded) vowel, /u/.

In a further effort to collect a set of data that would shed light on the full tone system of KTM, it seemed useful to include, not only smooth syllables (syllables ending in a vowel or a nasal consonant, and carrying one of the tones from Tone 1 to Tone 5) but also stopped syllables: syllables ending with a stop, and carrying tones 6 and 7 (the two tones of stopped syllables). In KTM, there are four possibilities for a final consonant in a stopped syllable, namely /-p/, /-t/, /-c/ and /-k/. My first choice was /-t/ because in my impression it seemed more frequent than others. However, a quick check in the word list showed that my initial impression was mistaken. Among 620 collected words, /-k/ was the most frequent final consonant (occurring in 44 words); the second most frequent was /-t/ (20 tokens), the third was /-p/ (13 tokens), and the fourth and last was /-c/ (6 tokens). Therefore, in terms of lexical frequency, /-k/ comes first. On the other hand, /-k/ has a strong identity in terms of coarticulation; in Vietnamese and Mường, this is reflected

in entrenched patterns of coarticulation within the rhyme. In Hanoi Vietnamese, the rounding of vowels /ɔ u o/ moves onto final velar consonants; the phenomenon only affects velars. In other words, among the six final consonants /-p, -t, -k, -m, -n, -ŋ/ of Hanoi Vietnamese, the velar consonants /k/ and /ŋ/ have special patterns of coarticulation. This is not only the case after rounded vowels. Haudricourt in his study of 1952 showed that Vietnamese has a phenomenon of shortening of vowels in front of velar consonants, and of palatalization of the final after a front vowel, such as the realization of /ɛk/ as [ɛc], and /ik/ as [ic] [48]. By my study, I found that /-k/ in KTM also tends to shorten a preceding vowel, so /-k/ was not selected for this experiment despite its status as the most commonly occurring final consonant.

In terms of formant movement, the final consonant that is likely to have the simplest coarticulatory influences is /p/: it lowers (slightly) the formants, especially at the transition from the vowel to the consonant, while /t/ has a stronger effect on F2 (which tends towards a target of roughly 1,800 Hz) which is especially salient after back vowels, for example in the sequence /ut/. In addition, /-p/, as a labial consonant, is the only consonant that does not conflict with the tongue position of the vowels. For those reasons, I chose /-p/ as a final consonant, hoping for data that would be easier to interpret.

To sum up, the final set of phonemes for creating target words consisted of: six initial consonants /p-, t-, k-, m-, n-, ŋ-/ , three nuclear vowels /i, a, u/, and final /-p/ for stopped syllables. The corpus of Experiment 2 is made up of the combinations of these phonemes and 5 + 2 tones (5 tones on smooth syllables and 2 tones on stopped syllables). Hence, we have: $[(6 \times 3) \times 5 + (6 \times 3 \times 1) \times 2] =$ a total of 126 syllables (90 tokens of smooth syllables, and 36 tokens of stopped syllables). Each syllable was said twice, hence the total number of tokens is $126 \times 2 = 252$.

An important issue when preparing the stimuli was the order in which tones are recorded. Explanations about tones are not the easiest part when using written stimuli, because the proposed system has to be learned entirely: to avoid interferences with Vietnamese, the orthographic representations used for Vietnamese tones are not used for KTM; the tone symbols chosen are numbers, which do not exist in any writing system with which the speakers are already familiar. Given the difficulty of cueing a tone, it seemed simpler at first to divide the data into seven sets, arranging them by tone: eliciting all the syllables that have Tone 1, then all that have Tone 2, and so on. This is easier for the consultants, as they keep tone constant while the syllable's initial and rhyme change. However, a drawback of this procedure is that the consultants' attention is drawn to the segments (consonants and vowels), which change from one syllable to the next, and are thus in focus, whereas (within each of the 7 tonal sets) tone is constant from one syllable to the next, and can therefore tend to become less and less clearly articulated, being, as it were, a less informational part of the syllable. Since the focus of the present research is on tone, what is wanted here is 'canonical' distinction: the fullest realization of the tonal templates. So it seemed advis-

able to change tone systematically from one token to the next; this solution was applied in the final set of recordings.

The target words were recorded inside the same carrier sentence as used in Experiment 1.

2.1.2.2 Training

The purpose of training was to enable the consultants to perform the experiment smoothly. The purpose of the experiment was not hidden from the consultants: I explained from the beginning that my study mainly focuses on the tones of KTM; and that the most important thing I expect from them is an accurate pronunciation of tones, so that the study can bring out the difference between the tones, and the characteristics of KTM tones. The consultants who participated in this study did not previously have a metalinguistic awareness of KTM tones: for instance, they did not know how many tones there were. On the other hand, they have an awareness of differences between Vietnamese and Mường tones, reflected in statements such as that KTM “does not have tones” (*“không có dấu”*), an observation by speaker M6 which I see as suggesting that the tone systems of the two languages are basically kept distinct.

Because KTM does not have a writing system, I borrowed Vietnamese writing system to write the vowels and consonants of the target words. I designed 126 cards corresponding to the 126 target syllables. Among these, smooth syllables were written in black letters on yellow cards, and stopped syllables in black letters on pink cards, so the consultant would distinguish these two sets more easily (see image on left of Figure 14).²

On cards, the target words were represented without tone: in other words, the tones were separated from the target words. They were put on a table beside illustrated photos (as shown in image on right hand-side of Figure 14). For example, the set for Tone 1 includes illustrations (photos) for: /kŋn₁/ ‘tree’, /t^haj₁/ ‘ear’, /βǎŋ₁/ ‘bamboo shoot’, and /t^hom₁/ ‘shrimp’, allowing the consultants to remember the correspondence between tone and number. This method avoids the use of the Vietnamese symbols for tone, which are familiar to the consultants but would entail high risks of carry-over of pronunciation from Vietnamese, no matter how much care is taken to emphasize that the intended tones are those of KTM and not of Vietnamese.

2. According to elementary color psychology as outlined on a website for a wide audience: <http://www.empower-yourself-with-color-psychology.com/meaning-of-colors.html>, yellow and pink are pleasant colors; these colors were chosen in the hope of helping reduce boredom and making consultants think of the experiment as a funny game. Specifically: being the lightest hue of the spectrum, yellow is uplifting and fun. Color psychology sees it as endowed with a potential to create enthusiasm and awaken greater confidence and optimism – properties that appear welcome to carry the consultants through a phonetic experiment. On this basis, I chose yellow for the most important part in the experiment: smooth syllables. As for pink, in color psychology, as in my own intuitive perception, it is a positive color inspiring comforting feelings, so it seemed an appropriate choice when moving from smooth syllables to closed syllables. Of course, this is not to claim that this choice is better than another, but simply to set out the reasons for selecting these colors.



Figure 14 – Experiment 2: Cards for (i) target words on the left-hand side (yellow cards for smooth syllables and pink cards for stopped syllables) and (ii) tones on the right-hand side.

The use of a separate indication for tone was also intended to direct the consultants' attention to tone: for subjects who have limited practice of writing, it appeared useful to draw attention to tone for each item, so as to guide them towards the intended tone.

The order of target words at elicitation (and hence, the order in which they appear in the recording) follows the following sequence: changing tones, then consonants, and finally vowels. For instance: /pi₁/, /pi₂/, /pi₃/, /pi₄/, /pi₅/, then /ti₁/, /ti₂/, /ti₃/, /ti₄/, /ti₅/, and so on. This order reflects the level of focus placed, in this study, on each part of the morphemes at issue: the most attention is directed to tones, and less is directed to consonants and vowels; among the latter, there is no considerable difference in terms of importance to the experiment, but in view of diachronic (panchronic) evolutions relating to tone, it seemed more plausible that interesting differences would appear between series of initials (stops vs. nasals) than between vowels, hence the choice to prioritize consonants over vowels, as it were.

Finally, the consultant was instructed to place each target word inside the carrier sentence which had been prepared earlier.

For the training, firstly, I explained to the consultants that this experiment purposed to gather all the tones of KTM, and it is important that they pay attention to the different tones and pronounce target words with the intended tone. Then, they were trained to distinguish the tones which appear, numbered from 1 to 7, on the poster, along with photos of example words (as on the right hand-side of Figure 14). Based on those photos, consultants recognized which tone was intended, and they combined it with the target words when the card representing the syllable was shown to them beside the tone. And of course, the target words with tone were spoken inside the carrier sentence just like in the first experiment. A photo of this experiment is shown in Figure 15.



Figure 15 – The training and performing of Experiment 2.

2.2 THE SPEAKERS

For performing this experiment, four native speakers were recorded: 3 male and 1 female. They are M1, M5, M6, and F1; the numbering is not continuous, because I use a numbering that refers to the full list of consultants who took part in recordings during my fieldwork on KTM. This full list is provided in Table 4.

Recruiting native speakers for recording was an important part of the data collection process. Because the recordings were conducted at field location, the elicitation sessions were relatively rapid and efficient: the speakers came by foot, or by a short bicycle or motorbike ride, on the day of recording, at a time that suited them; there was no issue of long journeys, or of taking speakers to an unfamiliar environment, as can happen when recordings are done far away from home. Another positive point is that the speakers' dialect backgrounds were fairly homogeneous: the consultants were chosen with the help of F1, who is familiar with them as members of the same rural community, and who selected proficient speakers, rather than the occasional newcomer to the community. As a result, all participants distinguished the five tones on smooth syllables and the two tones on stopped syllables that had been identified when eliciting the vocabulary list earlier on in fieldwork. This is an essential point, since the experimental setup had been devised for this tonal system, and cannot be applied if a consultant has a different system.

A further advantage of being based in the village is that it was not difficult to find consultants for recording. Instead of random selection of available subjects, we were able to take the initiative to select which people would be

Table 4 – The list of consultants, numbered sequentially according to the date of first participation in my fieldwork (starting in Feb. 2014).

numbering	name	date of birth	gender	code
1	Sa Thị Đính	08/02/1983	female	F1
2	Xa Thị Bích	15/02/1975	female	F2
3	Sa Thị Đang	13/7/1984	female	F3
4	Sa Thị Linh	14/12/1987	female	F4
5	Trần Thị Thắm	10/12/1961	female	F5
6	Hà Văn Chí	05/12/1979	male	M1
7	Sa Mạnh Hùng	25/8/1963	male	M2
8	Hà Văn Quyết	20/11/1977	male	M3
9	Đinh Văn Mối	04/01/1971	male	M4
10	Sa Văn Dẫn	05/11/1955	male	M5
11	Sa Mạnh Hồng	21/10/1971	male	M6

suitable for this recording, by communicating with them and understanding how they would relate to the task. F1 and myself took care to steer clear of consultants who would experiment trouble understanding the instructions; of consultants who are so eager to give advice that they become ‘experiment hijackers’, discussing the experiment instead of launching into it; and of consultants who have pronunciation problems that make them strongly atypical.

On the other hand, recording at field location also implied a certain number of difficulties. Among these, the greatest difficulty (in my experience) was to find male consultants for recordings: this proved much harder than finding female consultants. Because of agricultural activities, local residents are constantly busy with their business, especially men. The majority of men in the village are only at home twice a year, for the rice planting and rice harvest seasons. The remaining period of the year, they leave home to earn a living in the cities or to look after their family’s farm in the forest.³ Moreover, a few male speakers turned down the proposal to participate in linguistic fieldwork because they did not like the notion of being recorded or felt fear of ridicule and were not confident about taking part in the experiment.

Thanks to the kind help of local contacts (in particular F1 and M1), the data collection campaign could be carried to completion.

3. In Vietnamese, these forest farms are called as *tràng trại*: in addition to the main settlement, close to the paddy fields, families usually have a smaller farm higher up on the mountain, where they raise poultry and grow additional crops.



Figure 16 – Two recorders: (i) Fostex FR-2LE with two channels and (ii) Roland recorder R-44 with four channels.



Figure 17 – Two microphones: (i) AKG C535EB for table mic and (ii) Sennheiser HSP with MZA 900 P4 phantom power adapter for head mic.

2.3 THE EQUIPMENT

In term of recording equipment, I used two different steady-state digital recorders, one with two channels, and (later) one with four channels. On a first field trip, the equipment used was Fostex FR-2LE (shown in the left of Figure 16), a two-channel professional stereo recorder that allowed for the simultaneous recording of the audio signal in one channel and the electroglottographic signal in the other. A limitation is that only one audio channel can be recorded. By the time of the second field trip, an even better recorder was available, with four channels, allowing for the simultaneous recording of 3 audio channels and 1 electroglottographic channel: a Roland 4-channel recorder (in the right hand-side of Figure 16).

Two kinds of microphones were used: (i) microphone AKG C535EB,⁴ hereafter referred to for short as *table microphone* (see left-hand side of Figure 17), and (ii) two head-mounted microphones Sennheiser HSP4,⁵ with MZA 900 P4 phantom power (see right-hand side of Figure 17), hereafter referred to for short as *head mic*.

As the focus of the present study is on tone, it was essential to record an audio signal, as a matter of course; additionally, an electroglottographic signal was collected, as electroglottography is a noninvasive and portable device that offers some information about the evolution of vocal fold contact

4. Information of AKG C535EB microphone is available online: <http://www.akk.com/pro/p/c535eb>

5. Information of Sennheiser HSP4 microphone is available online: <http://fr-fr.sennheiser.com/microphones-serre-tete-hsp-4>

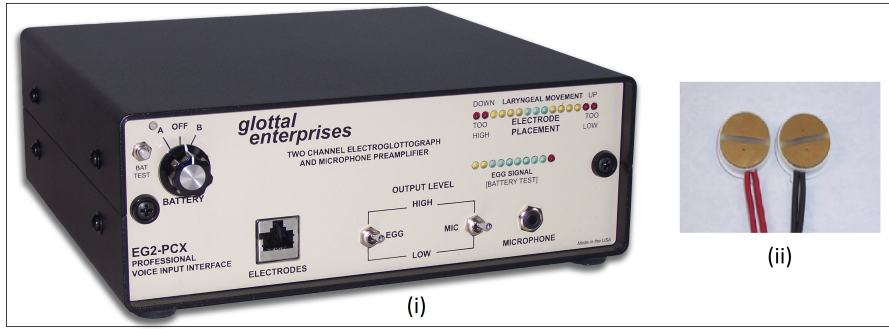


Figure 18 – The electroglottographic equipment: (i) Glottal enterprises EG2-PCX with (ii) dual-channel electrodes.

area over time. Electroglottography will be presented in more detail below, section 2.5; to summarize its usefulness in a few sentences: electroglottography is the ultimate touchstone for measuring fundamental frequency; it also provides indications on phonation types, which seemed promising to study glottalized tones in KTM. The model used is a Glottal Enterprises EG2-PCX⁶ (as shown on Figure 18). My experience confirms the promise made in the user manual: the Glottal Enterprises electroglottographs produce very low-noise EGG waveforms. The dual-channel configuration of the EG2-PCX2 enables it to provide an indication on electrode placement (too high, too low, or well-centered). Quantitative indication of vertical laryngeal movement was not used in the present study, where the dual-channel signal was recorded in *mono*, following the default configuration.

At recording, the position of these equipments was as shown on Figure 19. The two main equipments, the Roland recorder and Glottal enterprises EG2-PCX, were placed on the table. The dual-channel electrode collar was attached to the neck of the speaker to measure the translaryngeal electrical resistance at two adjacent locations; these two signals are combined as the main EGG output, which thus occupies only one channel in the resulting WAV file. Electrode placement was conducted as recommended by Nathalie Henrich, locating the last ring of the trachea and placing the electrodes above that point of reference. The input to the Roland steady-state recorder includes four channels, with four connection by XLR cables: (i) the first head mic, attached so that the microphone proper was above the corner of the mouth of the speaker; (ii) the table mic was put in front of the speaker and on the table (the distance between table mic and speaker was about 30-40cm); (iii) the EGG signal was recorded into the third channel of the Roland recorder; (iv) the second head mic (in a similar position with first one) was worn by the investigator-instructor (to keep track of the dialogue).⁷

6. Additional information about the Glottal enterprises EG2-PCX electroglottograph is available from the manufacturer's Web page: <http://www.glottal.com/Electroglottographs.html>

7. The second microphone was also used when eliciting conversations – not studied here –, in which case it was worn by the second consultant.

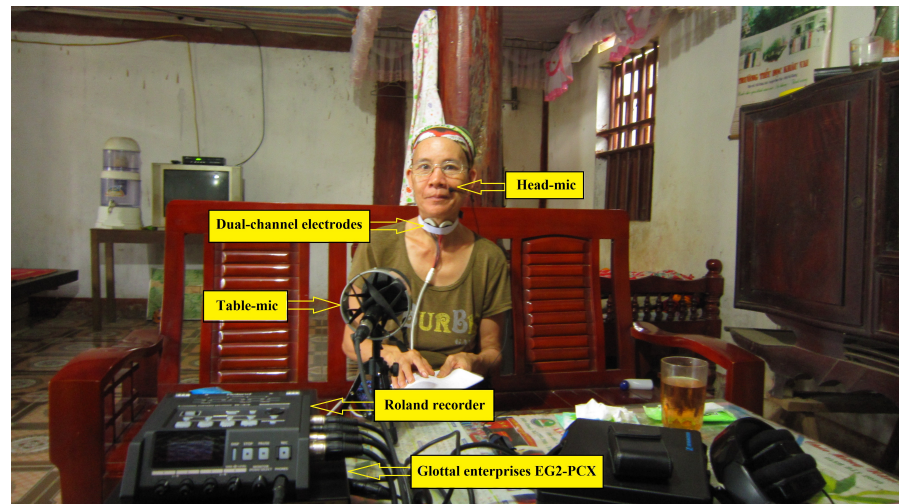


Figure 19 – The position of equipments on the recording (an illustration from the recording of speaker F5 in August 2015).



Figure 20 – Two support devices: (i) Uninterruptible Power Supply (UPS Santak TG1000) and (ii) LiOA SH 10000 voltage regulator.

Additionally, two supporting devices were used for power supply: an Uninterruptible Power Supply (UPS) and a voltage regulator were used to avoid the disruptive effects of power outages and large fluctuations in voltage. The models used were: Santak TG1000 UPS and LiOA SH-10000 voltage regulator, both shown in Figure 20.

But despite this double shield, a difficulty is the absence of grounding in the village. The electric network is not grounded. This is a serious issue in the case of audio recording. When using a head-worn microphone, there is conductive metal in contact with the speaker's skin during recordings, connected to the power grid; if the voltage differential between the two threads is not 220V vs. zero, but, say, 320V vs. 100V, there is a 100-V difference with the ground, which results in current down into the earth through the consultant's body. An internet search about electric shocks due to non-grounded microphones yields a harvest of horror stories about strong electric shock,



Figure 21 – A commonplace sight in Vietnam: transportation by motorbike.

with real threats to the consultants' peace of mind during recording and even to their health. This issue was encountered in the first fieldwork; Alexis Michaud reported encountering the same issue in remote areas of Yunnan. A good solution would have consisted in asking an electrician to provide grounding for the entire household, thus improving the family's security, in addition to solving the issue encountered in fieldwork. But no electrician was available locally to perform the operation. Given this situation, it appeared safest to rely on batteries instead, conducting all recordings on alkaline batteries.

The equipment was transported from Hanoi to the fieldwork location by motorbike, following local custom (as Figure 21).

The recordings are relatively good from a technical point of view, despite the fact that recordings were carried out on a farm, and not in a recording studio.

2.4 FILE EDITING AND ANNOTATION

At recording, both experiments were conducted one after the other as part of the same recording session. After the recording, all audio files (in WAV format) were split into two components corresponding to the two experiments. For instance, the data for speaker M1 was extracted from the original and renamed as:

- *crdo-MTQ_KTM_M1_TONEXP1.WAV*;
- *crdo-MTQ_KTM_M1_TONEXP2.WAV*.

These identifiers follow conventions used in the Pangloss Collection [88], an archive of language recordings in which the data is destined to be archived in due course. The identifier for each document begins with <crdo-> (the

former name of the archive) followed by the ISO code of the language in the Ethnologue catalogue of languages (the code for Mường is MTQ), then a short identifier of the document in capital letters (the underscore is allowed as a separator).

For this study, the data was from four speakers (M1, M5, M6 and F1), as mentioned above, but I mainly used the data of M1 and M6 because those are the best data with clear electroglottographic signals. The data of M5 contain numerous tokens of compromised signal, and hence were used mostly for qualitative analyses; and I completely set aside the data of F1 in the task of quantitative phonation-type estimation because of the weaker (more noisy) signal, a common situation for female speakers. The method used for analysis of the electroglottographic signal is especially demanding in terms of signal quality, as the estimation of phonation type requires identification of the glottis-opening instant, which can easily be drowned in noise (when the signal-to-noise ratio is not excellent). Therefore, the list of data provided as Table 5 only shows the audio files which were analyzed and used for this study. A full list of data is provided in the Appendix 8.

Using Roland 4-channel recorder yields files that are not in the stereo format which is most usual in digital audio (at the time of writing). Each audio files has four channels: (i) acoustic signal of native speaker from head mic, (ii) acoustic signal native speaker from table mic, (iii) EGG signal of native speaker, (iv) acoustic signal of instructor or cooperator. This can be handled without difficulty with Sound Forge, a professional digital audio editing software.

The bit-depth of the recordings is 24-bit, which allows for recording at a safe recording level, leaving headroom for sudden increases in voice intensity that may happen at any point during speech. Digital volume amplification was then conducted: the volume of the audio channel was normalized only once for each session, to preserve comparability across items within the same session. (The magnitude of amplification was on the order of 6 to 18 dB depending on sessions.) Needless to say, the original files were also retained for reference, without any processing after recording. Normalization was conducted using Sound Forge. The audio was not subjected to any other processing, such as filtering.

The annotation task consisted in indicating the temporal beginning and ending point of each syllable rhyme, for the target syllables and also for each syllable of the carrier sentence. Tone is non-segmental and it normally stretches over all of the syllable's rhyme part. Therefore, to measure glottal parameters relative to tone, one needs to annotate the rhyme part, exclude the initial consonant if it is voiced, one can annotate whole syllable if it is voiceless. Examples are shown in Figure 22.

During annotation, there were two special treatments, as illustrated on Figure 23.

(i) For tone analysis, a rhyme part should be have at least one cycle (i.e. the interval between two closing peaks on the derivative of the electroglottographic signal, as will be explained below). One cycle is enough for calculat-

number	name of audio file	speaker
1	crdo-MTQ_KTM_M1_TONEXP1_REPET	M1
2	crdo-MTQ_KTM_M1_TONEXP2_REPET	M1
3	crdo-MTQ_KTM_M5_TONEXP1	M5
4	crdo-MTQ_KTM_M5_TONEXP2	M5
5	crdo-MTQ_KTM_M6_TONEXP1	M6
6	crdo-MTQ_KTM_M6_TONEXP2	M6

Table 5 – List of recorded data used in this thesis: recordings conducted on latest field trip (Feb. 2016); speakers M1, M5 and M6; two experiments. “REPET” means this is the second time that the speaker participated in the experiment.

ing one F_0 value. However in a few cases the rhyme at issue is not present at all as a distinct temporal interval, so that it is not possible to assign it a beginning and endpoint – there is no identifiable portion of signal (not even one cycle) corresponding to the rhyme at issue. For this case, instead of a *Region*, I just place a *Marker* in SoundForge, with the mention “*note: syllable missing*”. The item is then deleted (manually) when exporting the Regions list from SoundForge for processing of the electroglottographic signal. (More details on the Regions list and how it is used further down the processing line will be presented in section 2.5.)

(ii) Because of coarticulation and hypoarticulation, there are a few cases in which a voiceless initial consonant has its manner of articulation modified: loss of closure for stops, realized phonetically as weak voiced fricatives (spirants in the sense of [84]). This phonetic phenomenon is called spirantization, a tendency that had a strong influence in the diachronic development of the Vietnamese consonant system [37]. This is not an issue for annotation: in that case, we just handle the rhyme the usual way, excluding the initial consonant.

These phenomena do not have a direct bearing on the target items: they only occur on frame syllables, not on the target syllable, which is generally hyperarticulated rather than hypoarticulated. For instance, issue (i) only occurred on the first syllable /**ja2**/ when the speaker tends to shorten the carrier sentence during a series of repetitions; and issue (ii) only occurred to the last syllable /**tǎŋ**/ because it is the end of sentence, so these syllables tend to be pronounced voiced.

At the same time as the beginning and endpoint of each rhyme was indicated, I also labeled every token. Labels are strings of numbers that consist in the concatenation of three part, referred to as A, B and C. Position A is the ordinal number of sentence which contains this syllable within the entire list. In fact, the value of this position depends on how many target syllables we have for each experiment. Therefore, position A runs from 01 to 10 at first experiment and from 001 to 126 at second experiment. Position B relates to

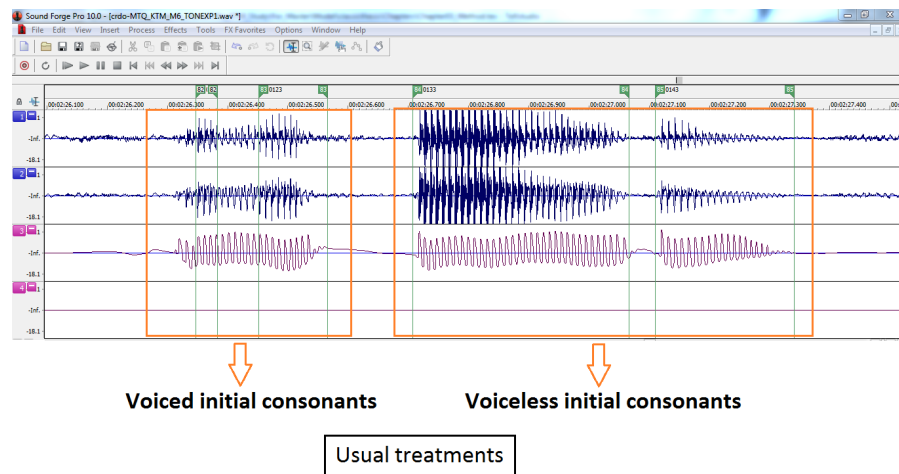


Figure 22 – Usual annotation: examples of different treatments for voiced and un-voiced initial consonants.

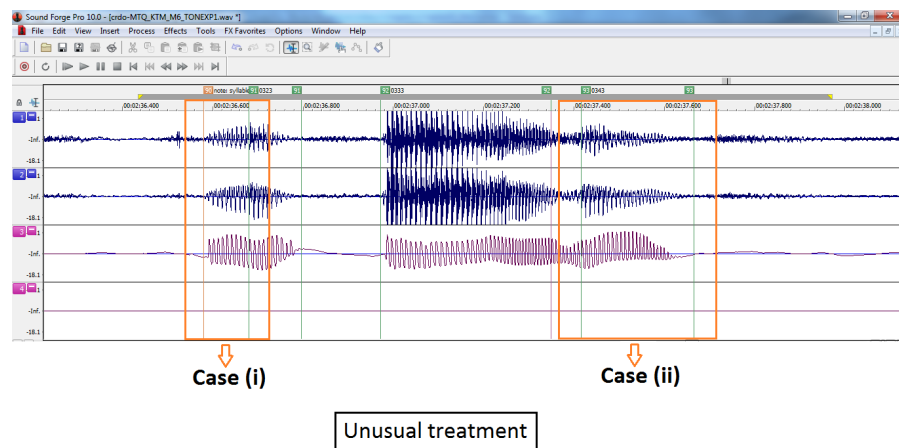


Figure 23 – Unusual annotation: examples of treatment for two special cases: (i) missing syllable, (ii) phonologically voiceless initial consonant realized phonetically as voiced.

Table 6 – Annotation scheme.

Position	Meaning	Number code
A	the order of sentence in experiment	01-10 at first experiment 001-126 at second experiment
B	the order of token inside sentence	1: /za2/ 2: /măt6/ 3: target word 4: /tǎŋ3/
C	repetition	1: first time 2: second time 3: third time

the position of the syllable rhyme within the sentence. Each sentence has four syllables, hence: value 1 means first token /**ja2**/ (2SG), value 2 means second token /**măt6**/ ‘to know’, value 3 means third token “*target token*”, and value 4 means fourth token /**tǎŋ3**/ (INTERROG). The last value in scheme, position C, indicates the repetition of item, value 1 for first time, value 2 for second time; for experiment 1 there is even a third repetition, indicated by value 3. For example, the indication 0111 means that this is the first syllable in the first sentence, at first repetition. A summary of the scheme is provided in Table 6. The precision of the annotation is key to precision in the quantitative values obtained at the next step: electroglottographic analysis, to which we now proceed.

2.5 ELECTROGLOTTOGRAPHIC ANALYSIS, USING A MATLAB SCRIPT BASED ON PEAK DETECTION FOR CALCULATING FUNDAMENTAL FREQUENCY AND GLOTTAL OPEN QUOTIENT

The method of analysis of the electroglottographic (henceforth, EGG) signal uses its derivative signal (henceforth, DEGG). This method is set out in full by Nathalie Henrich [55] and used (with some modifications) in [86, 89, 93], among others. The reason for adopting this method for the study of linguistic tone is because it holds promise for calculating parameters with a degree of replicability, precision and reliability that allows for comparison across vowels; across speakers; and eventually across languages. The present study about tone in KTM contributes to an ongoing trend to enrich the experimental setup in phonetic studies of tone in Vietic and in other languages of Vietnam, collecting not only audio but also other types of information.

2.5.1 Terminology: Some basic concepts

In term of methodology, the present study may be considered as an interdisciplinary study, to the extent that it combines linguistic insights with the use of recent technology. In particular, I used technology for linguistic processing to obtain quantitative results, in addition to qualitative observations. As this is a dissertation in Linguistics, intended for a readership of linguists, I do not assume prior familiarity with the technical aspects of phonetic analysis. This section sets out the background knowledge to the study, defining and explaining the key terms, some of which are likely to be unfamiliar to linguists.

2.5.1.1 *Electroglottography*

Electroglottography (EGG) was invented by Fabre in the mid-20th century. The initial report about the invention [33] was followed by further studies by the same author over several years [34–36]. Here is a simple explanation of the technique.

Electroglottography is a common, widespread technique that enables the investigation of vocal-fold contact area in phonation in an easy and noninvasive way. A high frequency modulated current ($F = 1$ MHz) is sent through the neck of the subject. Between the electrodes, electrical admittance varies with the vibratory movements of the vocal folds, increasing as the vocal folds increase in contact. [55, p. 1321]

In other words:

Though EGG does not provide an estimate of the air flow at the glottis, it is a precise (and noninvasive) measurement of the contact surface of the vocal folds, by means of a high-frequency, low-intensity current passed between two electrodes placed on the neck. [89]

EGG signal is the output. As it is a continuous signal, it can be stored in the same format as audio, and displayed with the same tools.

Henrich's paper goes into technical detail concerning the four main phases of a glottal cycle: (i) closing phase, (ii) closed phase, (iii) opening phase, and (iv) open phase. Increase in vocal fold contact area is reflected by the closing phase (itself followed by the closed phase) in the EGG signal; the moment of fastest increase in vocal fold contact area corresponds to the glottis-closure instant. Decrease in vocal fold contact area begins during the closed phase, and continues into the opening phase; the moment of fastest decrease in vocal fold contact area is the glottis-opening instant.

But the correspondence between these four main phases, on the one hand, and detectable events on the electroglottographic signal, on the other hand, is not easy to establish. Instead, the electroglottographic signal corresponding to one glottal cycle can be divided into two portions only, as shown on Figure 24. These two portions are named *closed phase* and *open phase* of

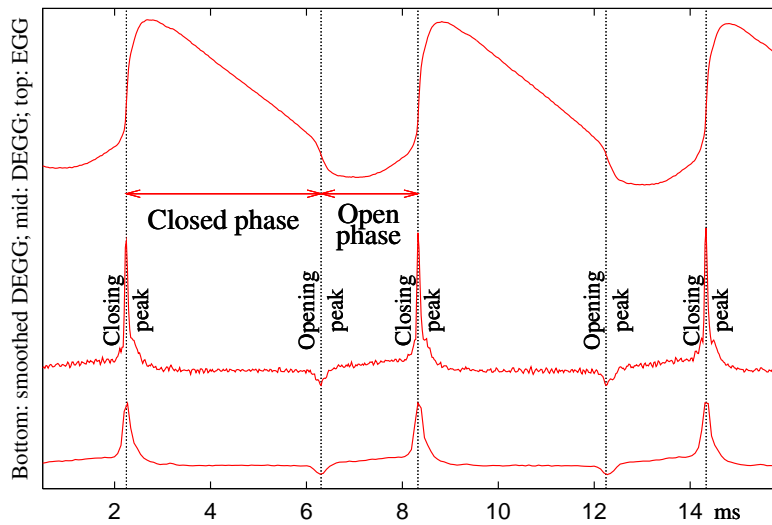


Figure 24 – Example of EGG and DEGG signals with indication of glottis closure and opening. Reproduced with permission.

the vocal-fold vibratory cycle, and defined as follows: the closed phase extends from a glottis-closure instant to the next glottis-opening instant; and the rest of the cycle is the open phase. (Greater detail is provided in [55, pp. 1321-1322], as well as [21], [24], [98].)

2.5.1.2 DEGG signal

DEGG signal refers to the *derivative* of the EGG signal which typically has a positive peak at glottis closure and a negative peak at glottis opening. Figure 24⁸ illustrates visually a synchronization of EGG and DEGG signals. In the case of clear signals, one closing peak is clearly visible for each cycle, corresponding to the peak increase in vocal fold contact area and considered as the beginning of the glottal closed phase, and one (less salient) opening peak, corresponding to a peak in the decrease in vocal fold contact area and considered as the beginning of the glottal open phase. Obviously, these peaks in DEGG signal relate to glottal opening and closing instants of EGG signal, hence they can be used to measure glottal parameters such as fundamental frequency and open quotient, the two parameters used in the present study.

2.5.1.3 F_0 and O_q

F_0 and O_q are two glottal parameters, which characterize vocal fold vibration.

F_0 is a shorthand notation for *fundamental frequency* (unit: Hz), the inverse of the glottal period (i.e. the inverse of glottal cycle duration). It is

8. This figure was reproduced (with permission) from reference [54], available from: <http://voiceresearch.free.fr/egg/index.html>

obtained by measuring the duration between two consecutive glottal closing instants, corresponding to a fundamental period; its inverse gives the fundamental frequency of the voice (the formula is simple: $F = 1 / T$). The values of F_0 have *pitch* as their perceptual counterpart: low F_0 is heard as low pitch, and high F_0 as high pitch.

O_q means glottal open quotient (unit: %). Measurement of O_q requires measurement of the duration of the glottal cycle, plus detection of the glottal opening instant. This allows for computing the glottis-open interval; the open quotient is the ratio of the open-glottis interval to the entire cycle (the ratio between open time and fundamental period) [89]. This can be stated as the following equation: $O_q = (\text{Open phase}) / (\text{Open phase} + \text{Closed phase})$. O_q is a parameter that relates to phonation types: low O_q demonstrates pressed phonation; medium O_q reflects modal phonation; and high O_q reflects flow phonation (whispery voice, shading into breathy voice). This relates to the following observation by Ladefoged and Gordon [43, 71]:

There might be a continuum of phonation types, defined in terms of the aperture between the arytenoid cartilages, ranging from voiceless (furthest apart), through breathy voiced, to regular, modal voicing, and then on through creaky voice to glottal closure (closest together). This continuum is depicted schematically in Fig. 25 ([43, p. 384]).

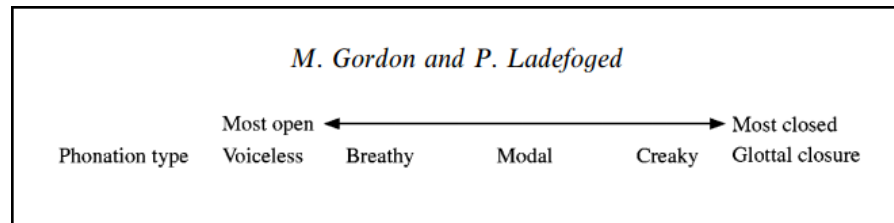


Figure 25 – Continuum of phonation types [43, p. 384].

2.5.1.4 *PEAKDET: an implementation of an analysis method based on Peak Detection*

There exist several scripts for analysis of the electroglottographic signal. That created by Nathalie Henrich [55] is designed for the singing voice; it is based on correlation, and is not designed to handle portions of voicing that are closed to the onset or offset of voicing. In speech, on the other hand, voicing is often interrupted and resumed. This difference in the input signals led linguists to write algorithms that are suitable for tracking glottal parameters in spoken language.

Peak Detection (henceforth *PEAKDET*) is a MATLAB⁹ script for calculating F_0 and O_q from electroglottographic signals. It is available online from the

9. MATLAB is a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis, numerical integration (Website: <http://www.mathworks.com/products/matlab/>)

website of *Tools for Electrolottographic Analysis: Software, Documentation and Databases* (link: <http://voiceresearch.free.fr/egg/index.html>). It is also available from COVAREP, a Cooperative Voice Analysis Repository for Speech Technologies (link: <http://covarep.github.io/covarep/>). COVAREP is an open-source repository of advanced speech processing algorithms hosted on GitHub: researchers in speech processing can store original implementations of published algorithms on COVAREP [27].

PEAKDET is designed for semi-automatic measurement: the results for each token are verified visually, and some parameters can be modified to adjust to the input signal. For instance:

PEAKDET shows the position of the first and last detected glottis-closure-instants on the electroglottographic signal and on its derivative, so that the user can appraise visually whether the full interval of voicing has been taken into account or not. This is useful in cases where the amplitude of the signal varies considerably within the portion of signal under analysis. (Excerpt from the PEAKDET user guide.)

2.5.2 *Step-by-step description about the process of analyzing electroglottographic signal*

I conducted electroglottographic analysis based on PEAKDET script. So in what follows, I am going to explain how I applied this script for my study. The diagram in Figure 26 summarizes the steps of the implementation of PEAKDET, dividing them into four basic steps: (i) setting default values and indicating the path to the WAV file to be analyzed; (ii) user verification of F_0 ; (iii) user verification of O_q ; (iv) saving output. These steps are described one by one in the paragraphs that follow.

2.5.2.1 *Step 1: Setting parameters and providing path to input files*

This step consists in setting parameters for the processing of the electroglottographic signal for one complete WAV file, hence, it has to be done once, at the beginning. When opening PEAKDET script, MATLAB will ask three parameters related to: (i) treatment of multiple closing peak; (ii) setting threshold for the detection of double peaks; (iii) choice of the number of data points for EGG smoothing. (Three command are displayed in MATLAB as Figure 27.)

In the case of my data:

- Variable #1: selecting 3 to use a value in-between (barycentre method) for peak detection in case of multiple closing peaks.
- Variable #2: setting 300 Hz for male and 500 Hz for female as a F_0 ceiling in the detection of double peaks.
- Variable #3: selecting one among the values from 1 to 3 for the number of points for DEGG smoothing, depending on the noise level of signal.

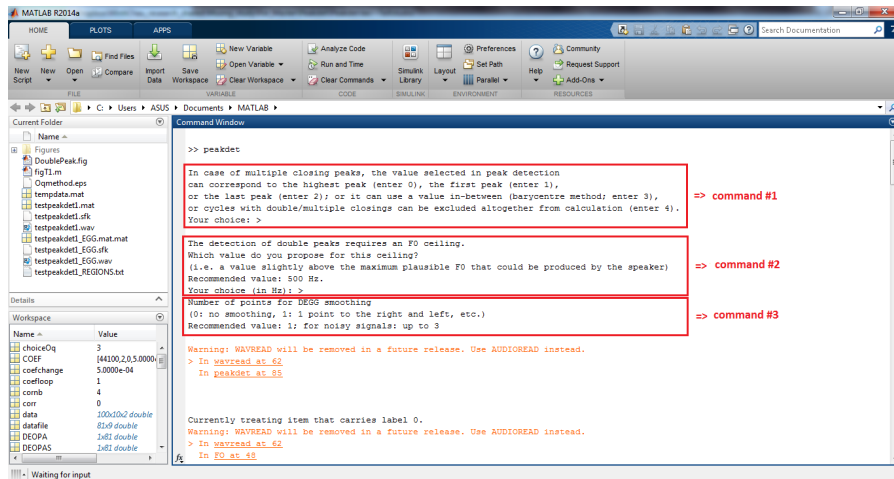


Figure 27 – MATLAB interface when giving three parameters: (i) treatment of multiple closing peak; (ii) setting threshold for the detection of double peaks; (iii) setting number of points for EGG smoothing.

saved as a text file in Notepad¹⁰. Figure 28 shows an example of REGIONS file in Notepad.

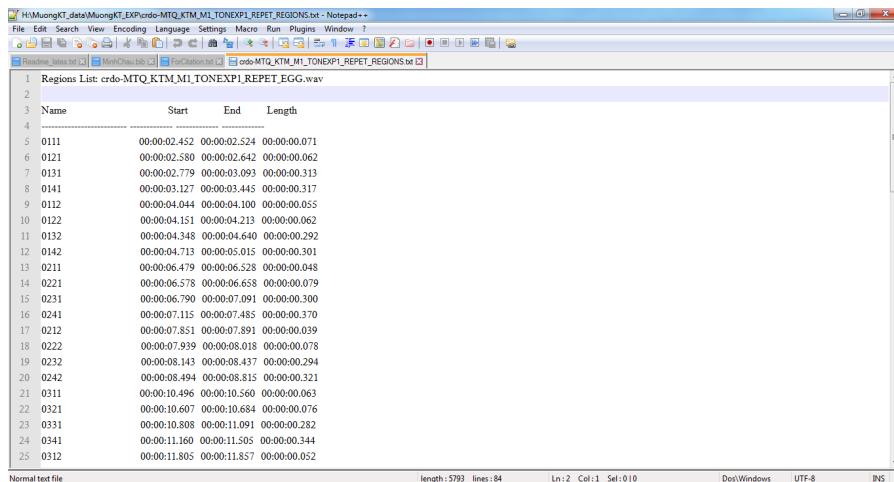


Figure 28 – An example of REGIONS file (.TXT) in Notepad. Experiment 1, Speaker M1

When PEAKDET has received the value of parameters and the path to the EGG and REGIONS files, it proceeds to the next step: peak detection on items, and display for user verification of F_0 and O_q .

2.5.2.2 Step 2: User verification of F_0

For each item, after smoothing signal and detecting peaks, MATLAB will plot three figures, as shown on Figure 29: (i) Result of analysis: Fundamen-

10. Notepad is a free text editor and source code editor that supports several languages, available from <https://notepad-plus-plus.org/>

tal frequency, and O_q calculated in four ways, (ii) Derivative of EGG (after smoothing), and (iii) Electroglottographic signal.

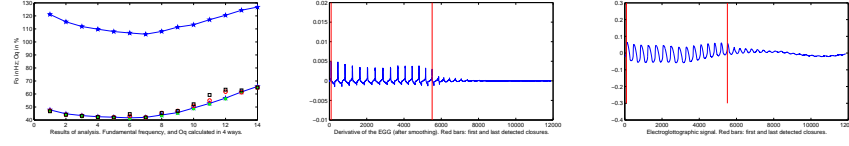


Figure 29 – Three figures obtained after smoothing signal and detecting peaks for one item: (i) Result of analysis: Fundamental frequency, and O_q calculated in four ways, (ii) Derivative of EGG (after smoothing), (iii) Electroglottographic signal.

At the same time that these three figures are drawn, MATLAB will show the values of F_0 as a listing and ask user verification in COMMAND WINDOW. We have four choices:

- If all the F_0 values are correct, type 0 (zero).
- If some of the periods went undetected, or extra periods were erroneously detected, enter 1 (one). You will then be asked to change the values of some of the settings.
- If you wish to correct some of the F_0 values, enter 2.
- If the coefficient is correct but the initial/final period(s) must be suppressed, enter 3.

The user's task consists in looking at figures and F_0 values, following the instructions of PEAKDET to make choices as appropriate in view of the results.

F_0 , as explained above, is computed from closing peaks because “*The fundamental period is given by the position of the first maximum, which corresponds to the time between two consecutive closing peaks*” [55, p. 1328]. In my study, closing peaks are mostly well-defined with a single peak, as illustrated on Figure 30. Hence, the curve of F_0 is usually continuous (as Figure 31). Therefore, in most cases we can quickly select first choice (“0” zero) for no correction and move on to the next step.

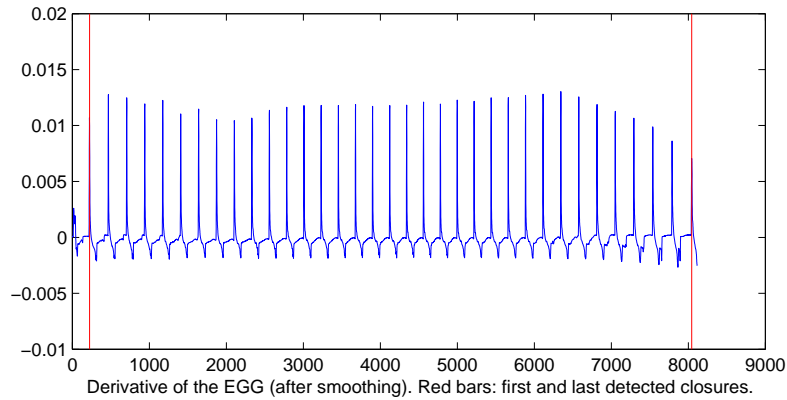


Figure 30 – An example of good DEGG signal with well-defined closing and open peaks.

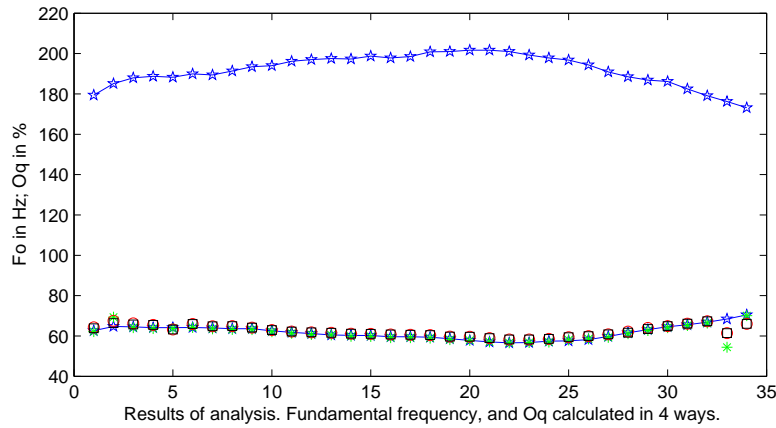


Figure 31 – Continuous F_0 and O_q curves as a result of good DEGG signal with well-defined closing and open peaks.

However in a few circumstances, we need some corrections because of non-continuous F_0 curve. This comes from two main reasons: (i) the first or last period(s) are out of the range found in the rest of the rhyme: they have widely different value(s) from the others; (ii) some of the periods went undetected, or extra periods were erroneously detected. The first case can easily be corrected by suppressing that/those period(s). For instance, item 2641 (data of experiment 2, speaker M1) showed a non-continuous F_0 toward the end at position of penultimate value (Figure 32) because PEAKDET detected two closing peaks after the rhyme had ended.

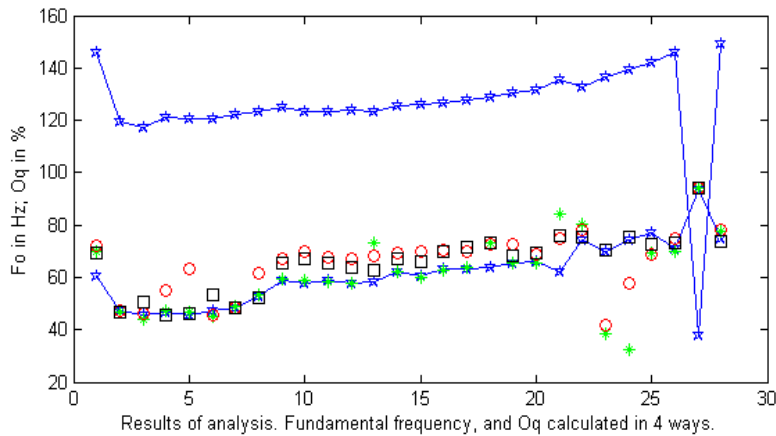


Figure 32 – First case of non-continuous F_0 : data of Experiment 2, Speaker M1, Item 2641.

This is not an artefact: the shape of the EGG signal and the derivative indicate clearly that two small closing peaks are present, i.e. the glottis closes at these points, and there is quasi-periodic fluctuation in vocal fold contact area. However, for the purpose of the present study, a crucial point is that these two small closing peaks take place after phonation has ended: the three symmetrical oscillations that precede the two closing peaks at the end

are typical of vestigial vocal fold oscillation at ‘soft’ (nonglottalized) offset of voicing; it would be obviously wrong to consider the time interval covered by these three small oscillations as constituting one glottal cycle. The corresponding F_0 value, at about 40 Hz (value before last in Figure 32) is clearly an artefact, and needs to be removed. Therefore, I suppressed the last two values.

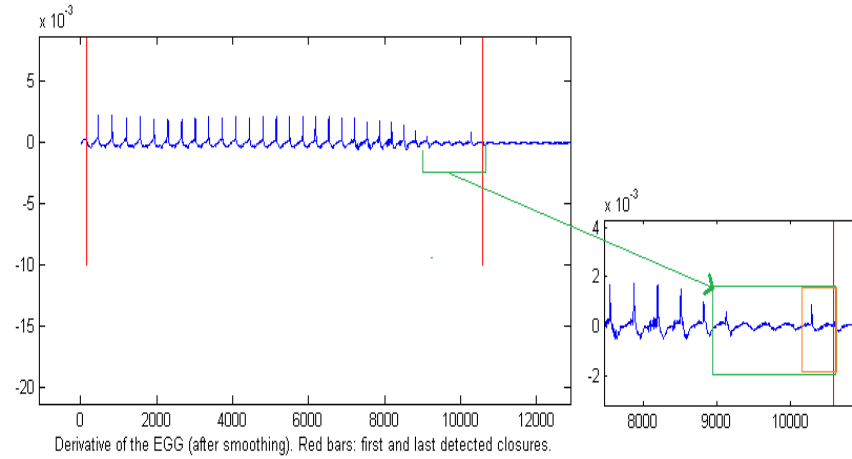


Figure 33 – DEGG of non-continuous F_0 , first case: data of Experiment 2, Speaker M1, Item 2641.

In the second case, it is necessary to adjust the value used as a threshold for the detection of closing peaks, so that PEAKDET can detect all periods (or leave out the spurious peaks, if extra periods were detected erroneously). To illustrate the case where a change in the threshold value is necessary because of undetected peaks, Figure 34 shows the location of undetected peaks in the DEGG signal, and Figure 35 shows the change in the resulting F_0 curve before and after threshold correction.

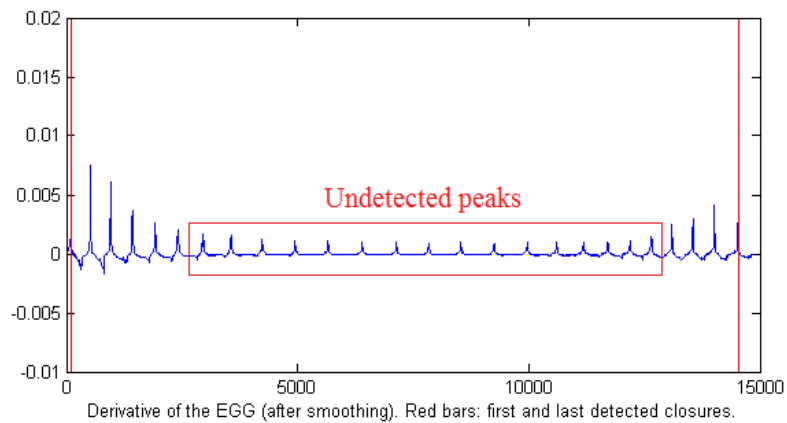


Figure 34 – An example of DEGG signal which shows undetected closing peaks.

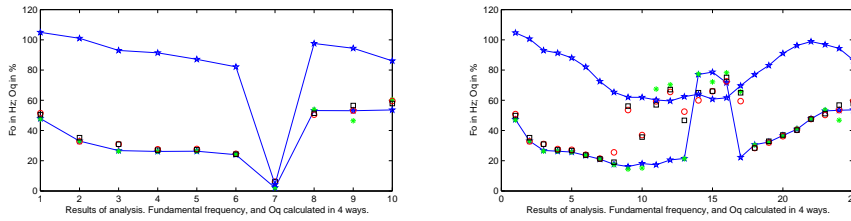


Figure 35 – The change of F_0 curve before (left) and after (right) threshold correction.

2.5.2.3 Step 3: User verification of O_q

Whereas F_0 is calculated solely on the basis of closing peaks, O_q requires the detection of opening peaks. PEAKDET will ask the verification of O_q after the verification of F_0 has finished. PEAKDET offers O_q calculated in four different ways: (i) maxima¹¹ on unsmoothed DEGG signal (in green), (ii) maxima on smoothed DEGG signal (in blue), (iii) barycentre of peak on unsmoothed DEGG signal (in red), (iv) barycentre of peak on smoothed DEGG signal (in black): see Figure 36.

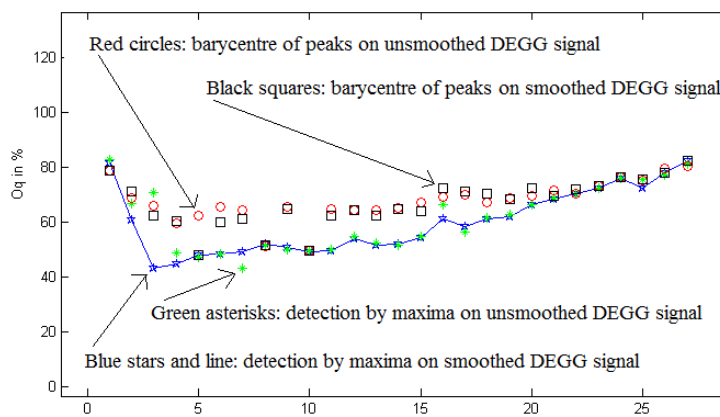


Figure 36 – O_q calculated in four different ways: (i) maxima on unsmoothed DEGG signal (in green), (ii) maxima on smoothed DEGG signal (in blue), (iii) barycentre of peak on unsmoothed DEGG signal (in red), (iv) barycentre of peak on smoothed DEGG signal (in black).

The methods are divided into two sets:

- Detection of the local minimum on the signal in-between two closure peaks; this method is applied twice : on the unsmoothed DEGG signal, and on the smoothed DEGG
- Analysis of the shape of opening peaks and calculation of a barycentre of the detected ‘peaks-within-the-peak’, giving each of the peaks a coefficient proportional to its amplitude. Again, this method is applied twice: on the unsmoothed DEGG signal,

11. Technically, ‘maxima’ here should be referred to as ‘minima’, since the peak is a negative peak.

and on the smoothed DEGG. ([54], online: <http://voiceresearch.free.fr/egg/index.html>)

Therefore, verification of O_q includes two steps: (i) choosing a method, and (ii) verification of the O_q results of that method. In simple cases, all open peaks of the item are well-defined peaks, which means that there is a single, well-defined negative peak inside each period, as in Figure 30. As a result, the four methods yield almost identical results of O_q , as shown by four coincident curves (as Figure 31). In this nonproblematic case, we can choose any method among the four methods, and there is no need for further correction of results before moving on to the next item.

However the simple cases were few. The majority of my data required more complex verification in the case of imprecise peaks in DEGG signal. ‘Imprecise peaks’ can mean several things. There can be several negative peaks during the opening phase, with none of them standing out as a main peak. Or there can be no well-marked opening peak at all.

Consequently, the result of four methods will be different, and the difference is often conspicuously visible by a mess of O_q curves as in Figure 37. In this case, we need a careful decision as to which method is appropriate, and which values among the set yielded by this method should be retained. By a verification of opening peaks in DEGG signal, we can choose a method which accurately reflects the well-defined peaks, then in next step the values corresponding to imprecise peaks should be suppressed. By a semi-automatic procedure, this verification is visual, and relative: decisions are left to the user’s appreciation. This does not appear unfeasible, but requires great care, and an investment of time. For instance, the analysis of total 1008 items of experiment 2 took over a week for careful examination and decision-taking.

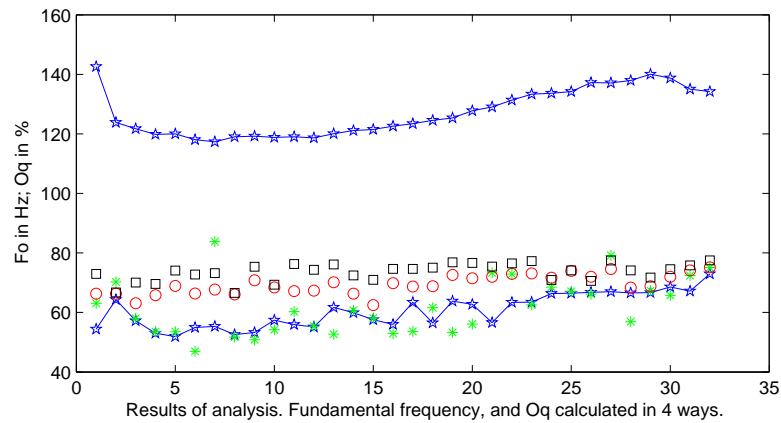


Figure 37 – An example of O_q curves with large differences across methods as a result of imprecise opening peaks. From data of speaker M1, experiment 2, item 1331.

Here is an example illustrating a case where there are imprecise peaks, and how those were dealt with. The example is from data of speaker M1, experiment 2, item 1331. After PEAKDET draws the three figures of (i) F_o & O_q ,

(ii) DEGG, and (iii) EGG, I zoom on the DEGG figure (38), taking a look to verify which peaks should be retained and which ones should be suppressed. Secondly, I consider the O_q values as plotted in Figure 39 to make a decision as to which is the best method in accurately reflecting the values of precise peaks. Finally, I start choosing the best method and suppressing values of imprecise peaks (those are set as zero, by convention) in COMMAND WINDOW. At that time PEAKDET will move on next item. In this example, the values in the middle are fine, and correspond to precise peaks, whereas the fit across methods is lost at both ends (beginning and end of the rhyme). In this situation, I decided to select method 1 (in blue)¹² and suppress values from 1 to 10 at the beginning and from 25 to 30 at the end.

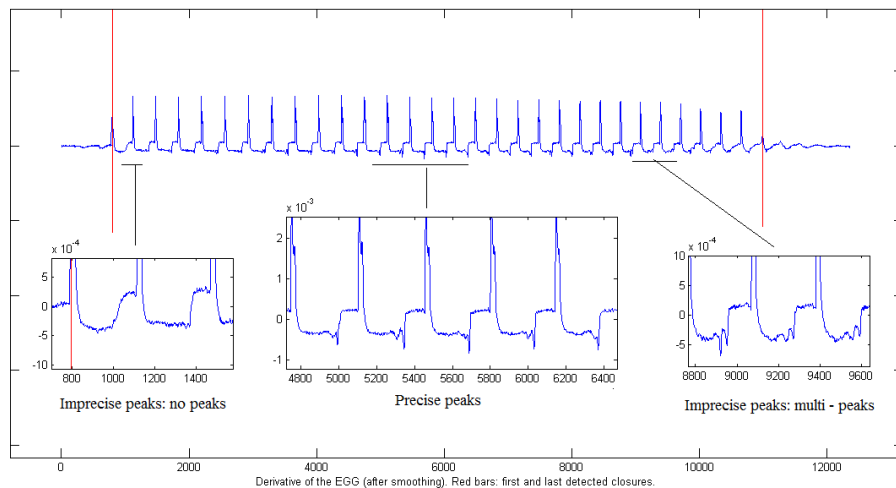


Figure 38 – An example of imprecise opening peaks: DEGG signal of item 1331, speaker M1, experiment 2. Abscissa: in samples (1 sample = 1/44,100 second).

Importantly, the values yielded by the four methods are all stored in the results matrix, and none is deleted during manual verification: instead, the values selected by the user are stored in a different column in the results matrix (the tenth and last column). It therefore remains possible, at any point in data processing further down the line, to compare the values chosen by the user to those yielded by fully automatic detection, or to compare the four methods. In other words, manual verification adds information to the results, and does not remove any.

After all the tokens have been thus inspected and processed, i.e. when F_0 and O_q have been calculated for all items in my data, PEAKDET saves the results as a “.mat” file containing all the variables in the work space: most importantly a matrix containing the results, but also other variables such as the one containing the labels of each token. The .mat file is then used by other Matlab scripts: for sorting into tone categories, for calculations, and for drawing figures. The end result is to offer quantitative and qualitative

12. The results shown in blue in Figure 39, obtained by peak detection on the smoothed signal.

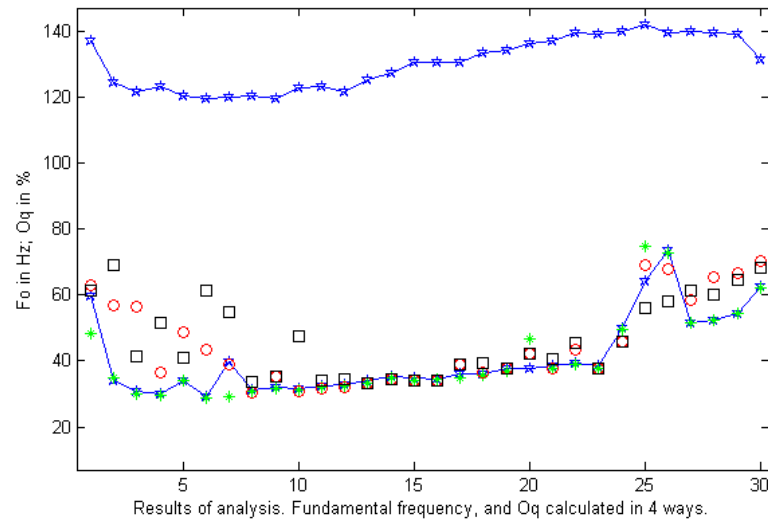


Figure 39 – An example of imprecise open peaks: O_q results of item 1331, speaker M1, experiment 2.

data for linguistic analysis of the tone system of KTM – the final purpose of the present research.

RESULTS

3.1 QUALITATIVE RESULTS: A MINIMAL SET EXEMPLIFYING THE FIVE TONES OF SMOOTH SYLLABLES

An overview of the tone system of KTM was presented in Chapter 1; this chapter will now go into a detailed description of each of the tones, starting with qualitative results.

The representation shown in Figures 40, 41 and 42 for Speakers M1, M5 and M6, respectively, were obtained from the analysis (through the procedure explained in the preceding chapter) of a minimal set illustrating the five tones over the syllable /paj/. (These data are from Experiment 1.) Inside each figure, the left-hand side represents F_0 curves, and the right-hand side represents the corresponding open quotient measurements. The average values of F_0 and O_q in Table 7 will be used as reference values to describe and evaluate all the results in this chapter. This value was obtained by averaging all values for syllable rhymes (target syllables and carrier sentences), as a rule-of-thumb reference value of the speaker's mean F_0 and O_q , for purposes of elementary normalization.

Table 7 – The average values of F_0 and O_q obtained by quantitative analysis of the EGG signal.

speaker	experiment	mean F_0 (Hz)	mean O_q (%)
M1	1	130	50
M1	2	140	49
M5	1	167	51
M6	1	112	48

3.1.1 *Speaker M1*

Tone 1, in green, is flat (phonetically level), in the middle part of the figure. It could seem as if it were right in the middle of the speaker's F_0 range: tone 1 is at about 120 Hz, in a phonetic space that covers a range from about 40 Hz to 200 Hz. On the other hand, the F_0 values at the bottom of the figure are those of Tone 4, a glottalized tone, and thus not directly indicative of the bottom of the speaker's F_0 range in phonation mechanism I ('chest voice'). Averaging over all the F_0 values extracted for syllable rhymes (target syllables, and syllables of the carrier sentence, i.e. a total of $5 * 4 = 20$ syllables), this speaker's average F_0 is 134 Hz. Seen in this light, the value of 120 Hz is

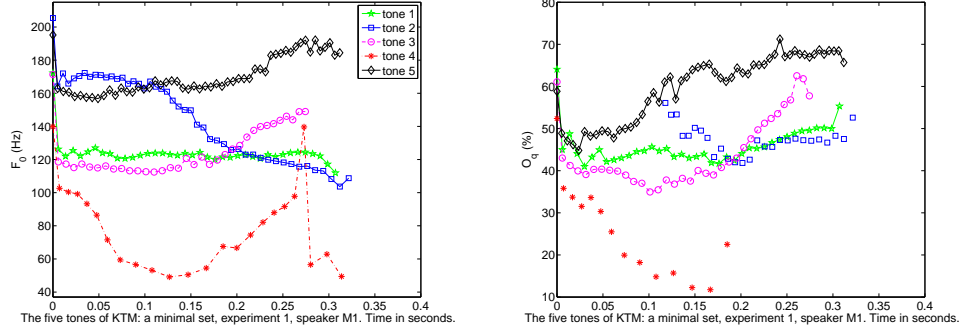


Figure 40 – F_0 and O_q measurement for a minimal set illustrating the five tones over the syllable /paj/, Speaker M1.

slightly below the speaker’s average. On the basis of this characteristic, I shall refer to this tone as a ‘low-mid’ tone.

The open quotient is in the mid part of the speaker’s range, slightly rising towards the end, an indication that the degree of vocal fold adduction decreases gradually, as expected towards the offset of modal phonation.

Tone 2 has a high beginning and drops to a low final F_0 value: as low as Tone 1, though not as low as the glottalized tone (Tone 4). The open quotient in the first third of the syllable cannot be estimated with any precision due to the absence of clearly distinct opening peaks on the derivative of the electroglottographic signal. On the portion where it can be tracked, the open quotient is in the middle of the speaker’s range; like for tone 1, the open quotient is slightly rising towards the end, indicative of modal/soft phonation at the offset of voicing.

Tone 3 is rising. The extent of the rise is not as great as that of the fall in Tone 2: from a minimum of 113 Hz to a maximum of 144 Hz, i.e. a rise of about 25%, whereas for Tone 2 the fall is from about 170 to 105 Hz, i.e. a decrease by over 60%. This is not surprising from a typological perspective, as a rise in F_0 is perceptually less salient than a fall (all other things being equal)¹. The open quotient rises sharply, which raises the issue of its correlation with F_0 . Open quotient and F_0 are not linked by a straightforward relationship of correlation: for instance, at the end of Tone 2, F_0 decreases while the open quotient tends to increase; and for Tone 1, the slight rise in open quotient towards the end does not correspond to an increase in F_0 : instead, it is interpreted as reflecting the gradual decrease in the degree of vocal fold adduction towards the end of phonation. In the case of Tone 3, the sharp rise in open quotient over the last third of the syllable corresponds to the rising part of the F_0 curve. It could be that, as F_0 rises, the phonation mechanism changes to mechanism II (‘head voice’) and the vibrating part of the vocal folds becomes smaller. This observation will be taken up again in the discussion of Tone 5.

1. On the asymmetry between a rise and a fall in F_0 , see e.g. [107].

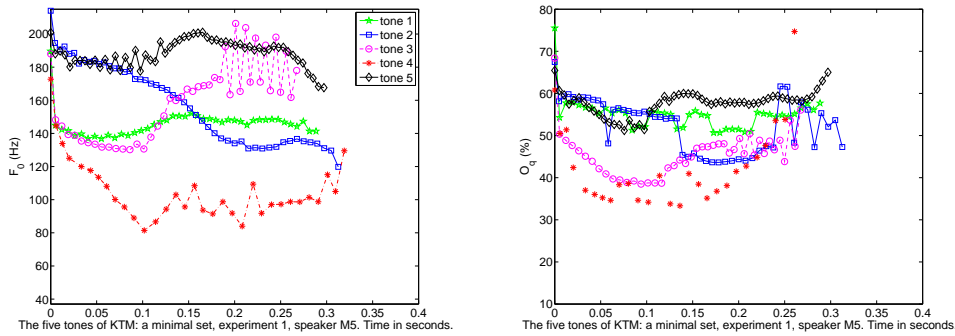


Figure 41 – F_0 and O_q measurement for a minimal set illustrating the five tones over the syllable /paj/, Speaker M5.

Tone 4 is glottalized early in the rhyme, though not quite at the beginning of the rhyme. The first two or three glottal cycles are in the bottom of the speaker's range, below the onset of both Tone 1 and Tone 3; from this low onset, F_0 then dips down in a telltale fashion, with a rapid acceleration, down to values that are only encountered in laryngealized voicing (about 50 Hz). F_0 then rises again, but the syllable ends in irregular phonation. The open quotient confirms this observation of a rapid transition from initial pressed voiced (open quotient below 40%) to laryngealized voicing. The vocal fold pulses still allow for an estimation of the opening instants and thus of the open quotient, which reaches values below 20%.

Finally, Tone 5 is rising, looking somewhat like Tone 3 but in a higher part of the speaker's F_0 range, and with a weaker rise, from 160 Hz to 190 Hz (+19%). Open quotient values are extremely high; like in the case of Tone 3 (which does not have such extreme values of open quotient), this raises the issue of whether a change in phonation mechanism accompanies the rise.

3.1.2 Speaker M5

Like for Speaker M1, Tone 1 is fairly level, but with a slight rise from an initial low value. This is one of the points that call for averaging over more tokens: this rise could be part of the tone's template; or it could just be a characteristic of this token that the speaker began at a F_0 level which he then judged too low, correcting through a slight rise. Tone 1 is, like for Speaker M1, slightly below the speaker's mean F_0 (averaging over the twenty syllables of the five target rhymes and their accompanying carrier sentences, the speaker's mean F_0 is 159 Hz). The open quotient is average – to the extent that it can be estimated with precision.

Tone 2 starts from F_0 values at the top of the range, and decreases within the first two thirds of the rhyme, to reach values which remain relatively stable over the last third. Open quotient decreases, with an apparent discontinuity; slight uncertainties about its measurement (some not-so-clear open-

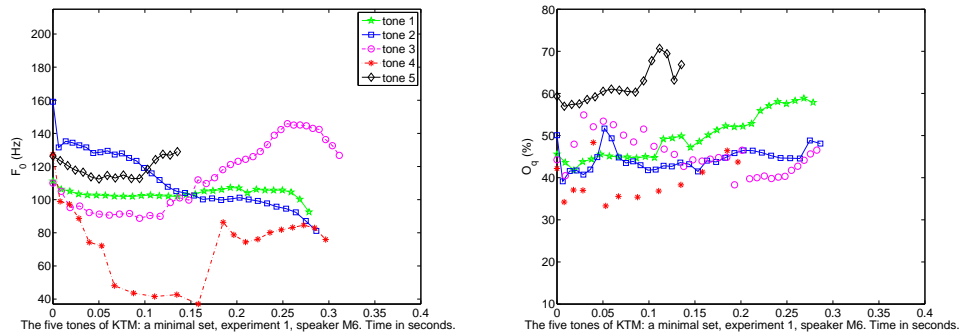


Figure 42 – F_0 and O_q measurement for a minimal set illustrating the five tones over the syllable /paj/, Speaker M6.

ing peaks) suggest that it is wise not to over-interpret these values. Here again, more data is necessary to advance towards an interpretation.

The rise in F_0 for Tone 3 begins earlier than for Speaker M1: before the middle of the rhyme. This rise is sudden, and reaches into the speaker's top F_0 range, with strong jitter. Interestingly, the open quotient during the first half of the rhyme is more similar to Tone 4 (the glottalized tone) than to the other three, which have distinctly higher open quotient.

Tone 4, labeled phonologically as a glottalized tone, does not display a noticeable lapse into creaky voice for this speaker. Neither the F_0 nor the open quotient reach as low values as in Speaker M1's example token for this tone. Still, the open quotient values below 40 are taken as clear evidence of constriction at the glottis (pressed voice).

Tone 5 is the highest tone, reaching a maximum of F_0 half-way through the rhyme. The open quotient is also the highest of all five tones, although not as spectacularly different from the other tones as in the sample set for Speaker M1.

3.1.3 Speaker M6

The mean F_0 for this speaker (calculated as above) is 113 Hz, hence a tonal space that is in a lower part of the figure than for Speakers M1 and M5.

Tone 1 is the most level of all five, with a moderate final dip corresponding to the offset of voicing. It is slightly below the speaker's average F_0 (mean value for Tone 1: 104 Hz). The glottal open quotient begins at values close to the speaker's average (averaging over all syllable rhymes: 47%), then rises to high values, though not quite at the top of the speaker's open quotient range.

Tone 2 is, like for M1 and M5, a high-falling tone. F_0 decreases from high values to the bottom of the speaker's range. Open quotient values are close to the speaker's average.

Tone 3 has a very pronounced rise. The open quotient has high values towards the beginning (though with some measurement uncertainties, due to

double opening peaks), which is the contrary of what is observed for Speaker M5; and it has lower values towards the end, which is the contrary of what was observed for Speaker M1. Our interpretation is that the defining phonological characteristic of this tone consists in the rise, and phonation type is left unspecified: there is no phonological specification to curb the influence of phonetic factors in the phonetic realization of the rising tone, such as the modification of phonation type by a transition to head voice (phonation mechanism II). If this view of things is valid, the expectation is that speakers will differ from one another in terms of patterns of habitual association of phonation type to Tone 3.

Tone 4 starts in the lowest part of the speaker's F_0 range: its onset F_0 is the lowest of the five tones. F_0 then decreases sharply, revealing early glottalization. Accordingly, the open quotient is lowest of all five tones. In the token chosen here, glottalization is relaxed towards the end of the syllable, returning to F_0 values suggestive of phonation mechanism I.

Tone 5 is much shorter than the other tones. The high open quotient values at the end of this tone (at the top of the speaker's range) suggest that vocal fold abduction takes place, interrupting voicing.

*A hypothesis:
different speakers
have different habits
in terms of phonation
types associated to
Tone 3.*

3.1.4 Summary, and expectations about results for larger data sets

In view of the above description, Tone 1 can be described as a mid-level tone, usually below the speaker's average F_0 . The relative flatness of F_0 curve is consistent across the three speakers. Hence, it still appears appropriate to characterize Tone 1 as having a *level F_0 curve*; this sets it apart from the four remaining tones, which all have more salient fluctuations in F_0 . The open quotient for Tone 1 is in the mid part of the speaker's range.

High-falling is the most typical realization of Tone 2. For all three speakers, this tone decreases from a really high beginning to lower final F_0 value. Its slope is uneven: more gradual at beginning and end, and more sudden in the middle. The open quotient is generally in the speaker's middle range, similar to Tone 1.

Tone 3 has a rise in F_0 , which can be considered as its distinctive characteristic, since none of the other tones shares this characteristic. The amplitude of the rise is not as great as that of the fall of Tone 2. In M5's data, there is jitter at the end of the F_0 curve of Tone 3, but this is specific to the token represented – serving as a reminder that qualitative examination of one token needs to be complemented by data from more occurrences.

The most remarkable feature of Tone 4, which distinguishes it from the four other tones, is glottalization. For all three speakers, both F_0 and open quotient are the lowest of the five tones, at the bottom of the speaker's range. The open quotient values dip below 40%, providing clear evidence of pressed phonation. In detail, the shapes of the F_0 and open quotient curves for the three speakers show salient differences across speakers.

The presence of higher open quotient on Tone 5 for all three speakers suggests that there may be habitual association of whispery phonation ('soft'

phonation) for this tone. The differences in shape of the F_0 curves between the three speakers suggest that what matters for this tone is that a high target be reached and strong falling F_0 be avoided. A final rise is an efficient way of achieving this goal (clearly distinguishing the curve from that of Tone 2, which is falling), but maybe not a necessary characteristic of the tone (a phonologically specified rising contour).

3.2 TOWARDS QUANTIFICATION: EXAMINING SETS OF SYLLABLES WITH THE SAME LEXICAL TONE

3.2.1 *Experiment 1: a small step towards quantitative results*

Experiment 1 is a small step towards quantification: verifying, on the basis of more data, the appropriateness of the characterization of tones proposed on a qualitative basis in section 3.1. This is only a small step because the data of Experiment 1 only consists of two minimal pairs. These pairs were repeated several times by the speakers, yielding several tokens of each tone. While repetition raises some methodological issues [117], it offers a convenient way of observing the degree of intra-speaker consistency in the realization of tones.

After the F_0 and O_q values of tones were obtained, as a matrix containing the values for all the syllables, various representations were attempted: tone by tone or in combination of all tones; with every single curve correspond to every single item or with a curve representing average value for all item. For average curves, there is also the choice to draw them with or without standard deviation, and in absolute F_0 values (in Hz) or in semitones. That yielded a large number of figures. This section only show a few figures that served for the analysis of result here. Extra figures are provided in an Appendix: chapter 10 of this volume.

Because the data of Speaker M1 has the best technical quality (clear signal and low noise) for both Experiments 1 and 2, this section will mainly consider and analyze M1's results, sometimes comparing with other speakers.

3.2.1.1 *Fundamental frequency (F_0)*

Figures 43 and 44 show three different views of the time course of F_0 for five tones by Speaker M1. Overall, they confirm the features of all five tones brought out through qualitative description in section 3.1. There are only slight differences with the qualitative results, such as that Tone 5 looks more level in the quantitative results. Its position at the top of the speaker's range is as expected on the basis of qualitative description. This indicates that contour should not be considered a defining criterion of this tone. The F_0 results confirm that the tones of KTM can be described as (i) two level tones distinguished by their overall register (Tone 1 towards the middle and Tone 5 towards the top of the speaker's range), (ii) a falling tone (Tone 2), (iii) a rising tone (Tone 3), and (iv) a glottalized tone (Tone 4). On the figures, the two level tones (Tone 1 and Tone 5) serve as two landmarks inside the tonal

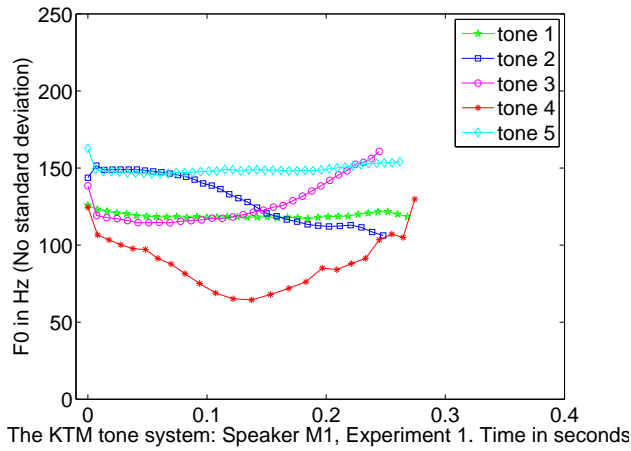


Figure 43 – The average curves of fundamental frequency (no standard deviation) for five KTM tones in the case of Speaker M1, Experiment 1.

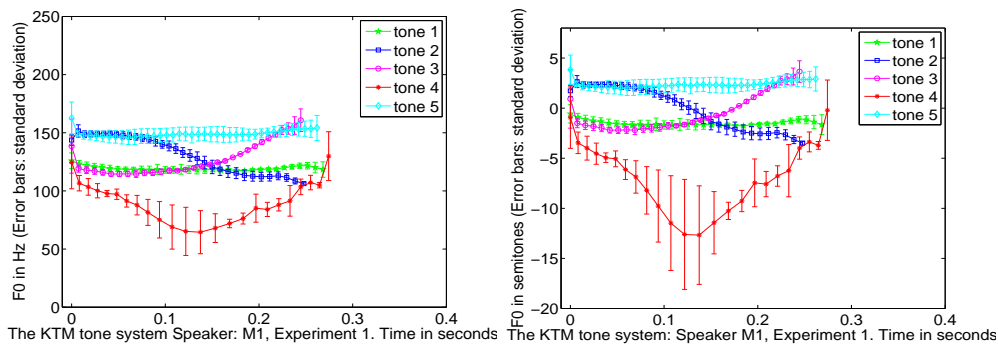


Figure 44 – The curves of fundamental frequency: (i) with only standard deviation (in the left); (ii) in semitones and with standard deviation (in the right) for five KTM tones in the case of Speaker M1, Experiment 1.

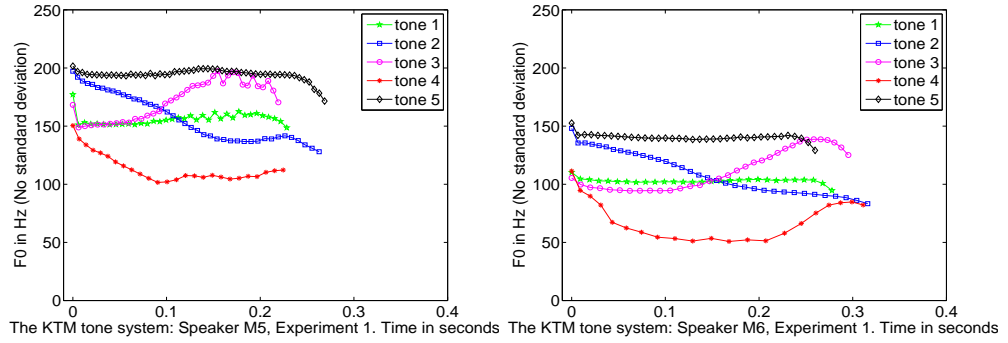


Figure 45 – Average F_0 curves of KTM tones system. Left: speaker M5. Right: speaker M6. No standard deviation. Experiment 1.

space, which can help to characterize the onset and offset of the other tones. Thus, Tone 2 starts in the top of range, the same position as Tone 5, then decreases and terminates near Tone 1, a ‘low-mid’ tone. By contrast, the beginning of Tone 3 almost coincides with the first part of Tone 1 before rising up to the area of Tone 5. Seen in this light, Tone 4 is a glottalized tone that reaches the bottom of the speaker’s range whereas its onset and offset are located close to the values found for Tone 1.

The tonal system (as viewed through F_0 tracings) thus appears relatively symmetrical, consisting of two parts. In the upper part, there are a low-mid tone contrasting with a top-high tone, and a falling tone (falling from the top to the low-mid of the speaker’s range) contrasting with a rising tone (rising from low-mid to the top of the speaker’s range). Meanwhile, in the lower part, the glottalized tone looks symmetrical along the time scale, with offset values close to onset values, and glottalization in-between. These observations also hold for data from the two other speakers (see Fig 45), suggesting that these are characteristics of the KTM tone system, not idiosyncratic properties (specific to one speaker).

Comparing across two graphs of Figure 44: on the Hertz scale (graph on the left of Figure 44), standard deviation is least for Tone 1 and Tone 3; it is slightly higher for Tone 2 and Tone 5; and it is unmistakably highest for Tone 4, especially in the glottalized part. The difference between Tone 4 and the remaining tones is brought out by the representation in semitones (right-hand side of Figure 44), recalculated on the basis of the speaker’s mean F_0 . Using this representation clearly sets apart the curve for Tone 4 from the others, highlighting the glottalization of Tone 4 which results in F_0 values that are both lower and more variable than for the other tones.

3.2.1.2 Glottal open quotient (O_q)

Figure 46 includes two graphs: the one on the left shows average O_q values without standard deviation and the one on the right shows the same data with an added indication of standard deviation.

The analysis of O_q raises some difficulties not found for F_0 , because a precise estimation of O_q requires the presence of a clear opening peak on the

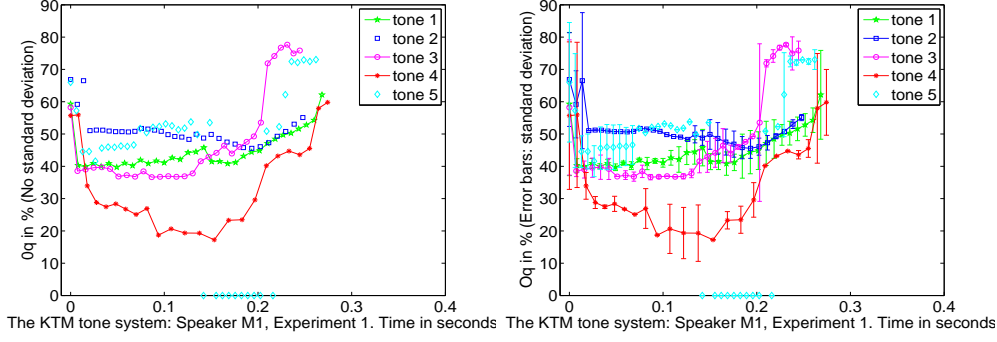


Figure 46 – Averaged O_q curves: (i) without standard deviation (left) and (ii) with standard deviation (right) for five KTM tones in the case of Speaker M1, Experiment 1.

DEGG signal inside each cycle, a condition that is not met in all cases, as discussed in Section 2.5. Technically, when visual inspection shows that opening peaks cannot be identified clearly, the O_q value for the cycle at issue is set at zero in the results matrix, and excluded from the calculation of averaged curves.

In particular, open quotient values are missing midway through the rhyme in *all* tokens of Tone 4. This is reflected in Figure 47 as an averaged curve that is discontinuous, plunging to zero for seven cycles. A return to the DEGG signals, exemplified in Fig. 48, reveals that it is not coincidental: there is a lapse into creaky voice at this point in the rhyme in all six tokens, with complex-repetitive patterns that do not allow for the calculation of the glottal open quotient.

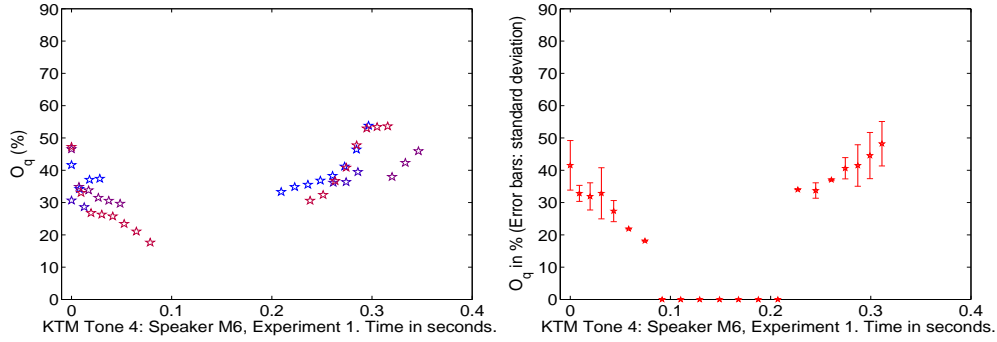


Figure 47 – An example of glottalized Tone 4 with the open quotient curves missing values midway in the case of Speaker M6, Experiment 1: (left) Individual items: one curve per item, (right) Average data with standard deviation.

This difficulty does not exclusively occur in Tone 4 (because of complex-repetitive patterns). There are also some cases for some tokens of other tones. This is a relatively common situation. I have calculated the proportion of excluded values for each speaker: see Table 8.

There are salient differences across the speakers and also across experiments, but the figures are similar with those obtained for an unrelated lan-

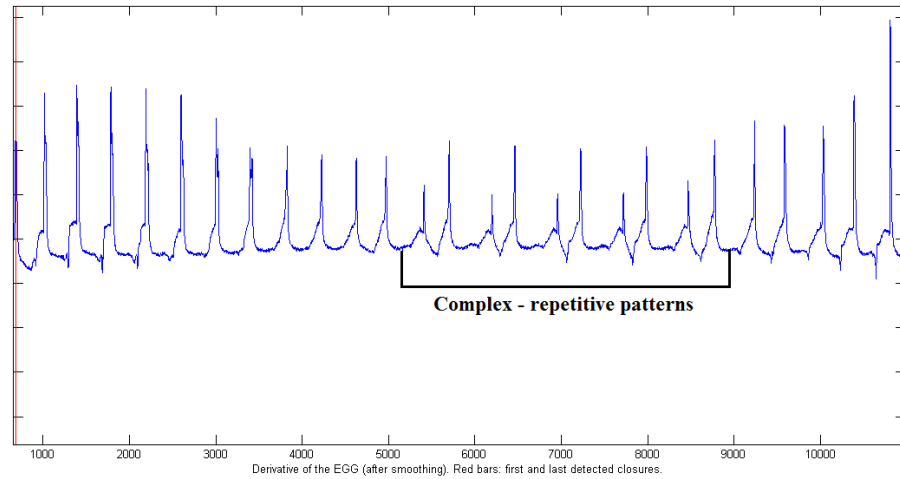


Figure 48 – An example of DEGG signal, with complex-repetitive patterns that do not allow for the calculation of the glottal open quotient. Speaker M1, Tone 4.

Table 8 – The proportion of cycles for which O_q could not be estimated on the basis of the EGG signal.

Experiment	Speaker	O_q Missing (%)
1	1	34
1	5	9
1	6	8
2	1	19

guage (Tamang, Sino-Tibetan family) [86, p. 235]. In that study, the proportion of excluded values of Speaker M2 is only 3.47% whereas the proportion comes up to 32.8% for speaker M4. The following reasons may play a role:

1. State of speaker: vibratory movements of the vocal folds are influenced by health status, state of mind, amount of alcohol consumed, etc.
2. Recording device: electrode placement affects directly the EGG signal. If two electrodes are placed too high or too low on the neck, too tight or too loose, the signal can easily become so noisy that opening peaks get drowned in noise.
3. Investigator decisions: as mentioned in Section 2.5, PEAKDET is designed for semi-automatic measurement: the results for each token are verified visually. Some users may have higher standards than others: when there are plentiful recordings, the phonetician may choose to exclude some not-so-clear-cut cases which would be retained in a study based on less abundant materials. Therefore, the results are unavoidably subjective to some extent.

Keeping in mind this inherent limitation of O_q as a parameter – that it cannot be meaningfully calculated in cases where there is no clear, well-

defined opening peak –, those open quotient values that could be calculated are still useful. The average of O_q of speaker M1 is 50% (as shown in Table 7). It can be seen from Figure 46 that most tones have open quotient value in the mid part of the speaker's range. In particular, Tone 2 and Tone 5 have similar values at about 45-50%. Tone 1 and Tone 3 are slightly lower, at about 40-45%. Tone 4 is remarkably different in comparison with the four other tones. Its open quotient switches rapidly from initial pressed voiced (open quotient below 30%) to laryngealized voicing, which reaches values below 20%, i.e. the very bottom of range. These characteristics are similar in M6, whereas O_q values of M5 are mostly restricted to the range from 40% to 60%. Thus it can be easily noticed that F_0 and O_q for Tone 4 are both very different from the other tones, for the same reason: glottalization is strongly reflected in these two parameters, with longer cycles, and longer closed phase inside each cycle. In other words, we have a good tool to detect glottalization, as information from O_q (when available) and from F_0 confirm each other.

Standard deviation of open quotient provides indirect evidence about missing values in O_q curves. Figure 46 shows the standard deviation in open quotient of Speaker M1. Due to some missing values (as mentioned above), Tone 5 has a discontinuous curve with 10 values set to zero. All five tones have several values which are based on one data point only, with the result that there is no standard deviation. The reason is that most values have been suppressed because of imprecise opening peaks. A high standard deviation of open quotient, together with a jigsaw-shaped averaged curve, provide indirect evidence that we cannot have an accurate evaluation of the tone's open quotient.

As we known, open quotient is a parameter which relates to phonation types. Values above average reflect 'relaxed', 'flow' phonation; values below that reflect more 'pressed' phonation. On the basis of figures and description of O_q above, some phonation features of KTM tone system can be suggested here. There is a two-way phonation distinction among five tones: modal and glottal. Four tones have modal phonation with O_q values around the middle of range: Tone 1, Tone 2, Tone 3 and Tone 5. Tone 4 is characterized by glottalized voice quality, corresponding to the lowest O_q : at the bottom of range.

3.2.2 Experiment 2: exploring intrinsic and co-intrinsic effects

Experiments 1 and 2 are homogeneous in many respects:

1. Same speakers: M1, M5 and M6.
2. Same environment: the same place (at the house of consultant M1) and the same time (always performing two experiments in the same recording session).
3. Same recording equipment (as mentioned in Section 2.3)
4. Similar method: using images to remind the speaker about which target syllable is intended, then requiring them to say it inside a carrier sentence.

The main difference between the two experiments is the amount of data. Experiment 1 has only two minimal sets which are repeated twice (i.e. 4 tokens for a tone), whereas Experiment 2 was set up with 18 minimal sets which are moreover repeated twice (means 36 tokens for a tone). This significant disproportion makes comparison of results almost impossible, hence Experiment 1 is essentially considered as a “stepping stone” to provide the basis for Experiment 2.

Like Vietnamese, KTM has two contrasting tones on stopped syllables, thus we can say that KTM is a dialect with a tone system of 5+2 tones.² With a view to obtaining a full picture, these two tones are also included in Experiment 2 and analyzed as independent tones.

Like for Experiment 1, the results for Experiment 2 are presented tone by tone, reporting on the F_0 and O_q of the 5+2 tones. Among other differences between the two experiments, the average value of F_0 (averaging across all syllable rhymes, including the carrier sentence) for the same speaker is not identical: M1’s average F_0 value was 130 Hz in Experiment 1; in Experiment 2, it rose to 140. Higher pitch usually goes together with more implication and more attention. This suggests that Speaker M1 was more careful (or even somewhat stressed) when going through Experiment 2. This experiment is significantly more challenging than Experiment 1 because the speaker has to produce KTM tones on minimal sets of non-words, using a phonological pathway to arrive at the intended syllable (putting the syllable together on the basis of its constituting elements), and not a lexical pathway (identification by means of a picture).

In Experiment 2, there occurred an accident involving Tone 1 and Tone 2. Figure 49³ shows that Tone 1 and Tone 2 are confused with each other. Results of both F_0 and O_q overlap to such an extent that they are not distinguishable. Comparison of F_0 curves with those obtained in Experiment 1 suggests that the confusion in experiment 2 consisted in realizations of Tone 1 as Tone 2: the F_0 curves are falling with a big variation in onset (from 150 Hz to 200 Hz) and lesser in offset (from 110 Hz to 120 Hz), looking more like canonical realizations of Tone 2 than of Tone 1.

O_q values, shown in Figure 50, do not provide any useful indication here, because Experiment 1 revealed that open quotient of Tone 1 and Tone 2 are

2. Note, however, that it is possible to reanalyze the two tones of stopped syllables as allotones of two of the ‘smooth’ tones. Such is the choice made for Vietnamese by Marc Brunelle and James Kirby: “...Northern Vietnamese, which has six contrastive tones in open syllables and two contrastive tones on checked syllables, can be analysed as having six or eight tones. Here, we assume that the two checked tones can be analysed as allotones of two of the open tones and settle for a six tones analysis.” [17, p. 85] (The page number refers to the preprint version available from: <http://www.lel.ed.ac.uk/~jkirby/docs/brunelle2015reassessing-preprint.pdf>)

3. In this figure, the different color of lines change from red to blue to reflect the order of tokens in experiment. Red curves means the beginning of set, maroon curves means the middle of set, and blue curves means the end of set. This color code is also applied for all figures which show one curve for each item. The aim is (i) to make the lines easier to tell apart from one another and (ii) to provide a visual indication of changes in the speakers’s behaviour in the course of the recording (commonly observed changes include gradually decreasing F_0 : see [95])

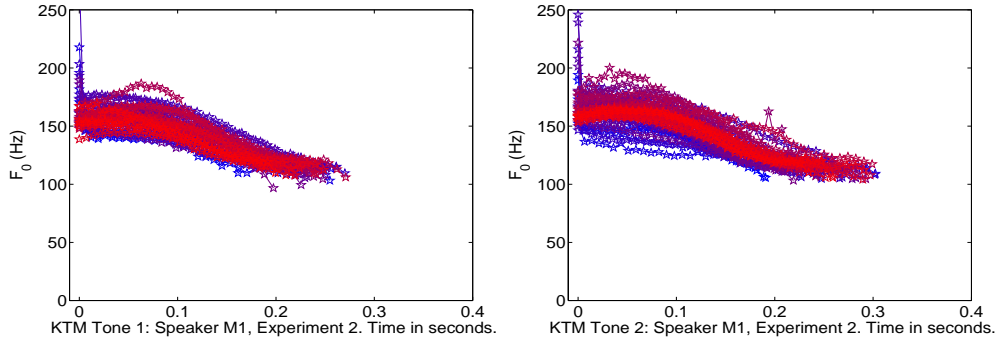


Figure 49 – Curves of F_0 for Tone 1 (left) and Tone 2 (right), in a data set (speaker M1, experiment 2) where these two tones were not distinguished consistently. (Individual items: one curve per item).

close, both in the middle of range: the values observed in Figure 50 (in the range from 40 to 60%) are compatible with interpretation as either Tone 1 or Tone 2.

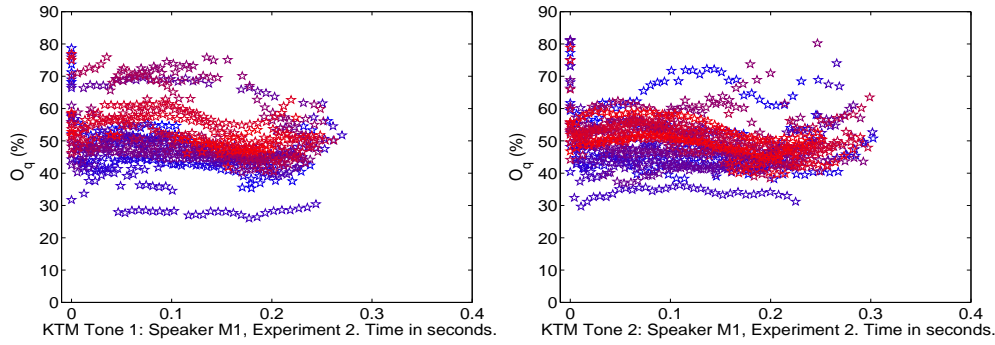


Figure 50 – O_q curves suggesting a confusion of Tone 1 and Tone 2 by speaker M1 during recording. Left: Tone 1, right: Tone 2. Experiment 2. Individual items: one curve per item.

The same confusion of Tone 1 and Tone 2 also happened with other speakers, so it can be considered as a bias of the experiment. I may have been too demanding with the speakers in this experiment, providing insufficient preparation. While I assume full responsibility for this glitch, and will learn from it to design more appropriate experimental setups in future, it may nonetheless be revealing about the KTM tone system. Why would speakers confuse Tone 1 and Tone 2, rather than another pair? The tonal spaces of these two tones are really close: small difference at the beginning (on the order of 25 Hz) and even smaller difference towards the end, where F_0 values overlap a lot. In addition, they both have modal phonation with similar open quotient. Moreover, there is probably an influence from Vietnamese – a language whose influence is encountered at various levels when studying Mường. Tone 1 of KTM has no close equivalent among Vietnamese tones whereas Tone 2 similar to Vietnamese tone A2, “*huyền*”. As a result, speakers of KTM tend to identify both KTM Tone 1 and KTM Tone 2 with

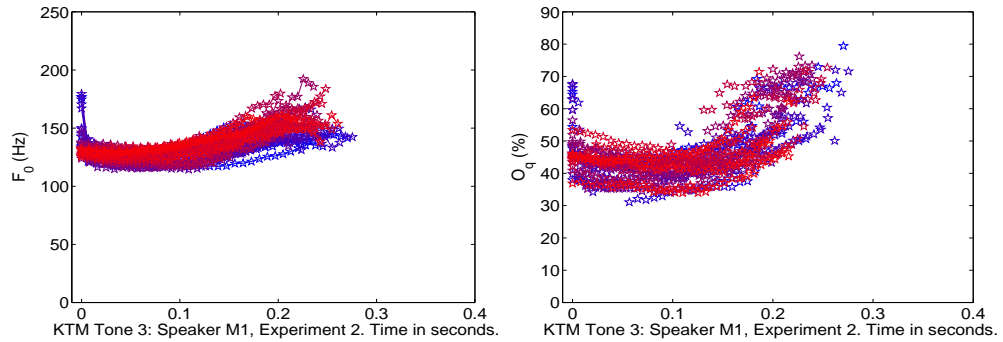


Figure 51 – The curves of F_0 and O_q for Tone 3, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right (Individual items: one curve per item).

Vietnamese tone A2. This experiment perhaps made speakers think back to spelling lessons at school (i.e. in Vietnamese).

Now turning to Tone 3, the results of Experiment 2 confirm its most salient characteristic: Tone 3 is a rising tone. The left-hand part of Figure 51) brings out an interesting contrast between the beginning and end of the rhyme in terms of dispersion of F_0 values. At the beginning of the rhyme, there is high consistency. The curve has a stable part in its first half, preceding the clearly rising part (second half); the second half is more variable than the first, as if realizations fell into two groups: either (i) rising until the end, or (ii) rising then stable again or even slightly going down towards the end. Given this situation, it seemed prudent to investigate the hypothesis whereby one of the two patterns reflects the usual realization of Tone 3 in KTM and the other is influenced by Vietnamese Tone B1. If referring to stylized F_0 curves for the Vietnamese tone system proposed by James Kirby [61, p. 386], we will recognize that the first pattern (rising throughout) is similar to tone B1 in Vietnamese. By contrast, the second pattern is perhaps more typical of KTM Tone 3. (This is simply a hypothesis, to be investigated further.)

The open quotient rises sharply towards the end of the rhyme, concomitant with the rising part of the F_0 curves. This tone has the most stable O_q values in comparison to all tones. The interpretation of this rise in O_q is not straightforward. Since it goes hand in hand with a rise in F_0 , some of the rise in O_q may be a low-level phonetic consequence of the change in F_0 . This issue will be taken up in the Discussion.

Tone 4 has a complex F_0 graph in the left of Figure 52: the middle part of the curve has saw-like jitter from one cycle to the next.

This tone is definitely glottalized with both F_0 and O_q the lowest of five tones, in the bottom of the speaker's range. Besides, another characteristic of Tone 4 as realized in the context of the two experiments reported here is complex-repetitive patterns on DEGG signal, a telltale indication of creaky voice (Figure 53 is one among many examples). Complex-repetitive patterns create a difficulty for O_q calculation, because one short cycle with relatively high O_q can alternate with a longer cycle with low O_q . As a result, open quotient shows some unusual values reaching up to 80%, within curves that are

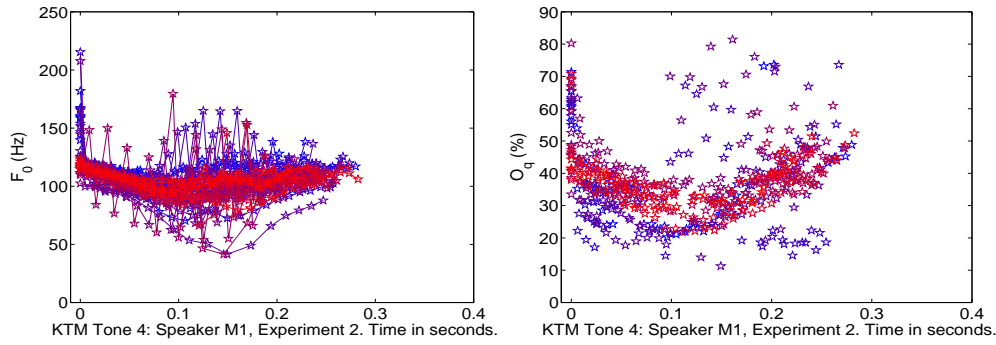


Figure 52 – The curves of F_0 and O_q for Tone 4, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right (Individual items: one curve per item).

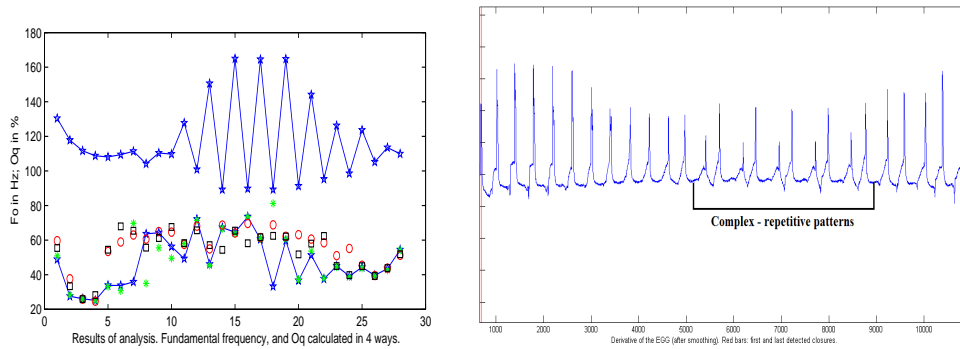


Figure 53 – An example of creaky Tone 4 with a curve of fluctuating F_0 and corresponding DEGG signals in the cases of Speaker M1, Experiment 2, item 2931.

overall below 40%. In addition, it can happen in Tone 4 that the density of O_q values is much lower than for the other tones because of imprecise opening peaks (in which case the O_q values are set at zero during verification, as explained in the Method section). Those constitute evidence for analysis as a glottalized tone with constriction at the glottis and sometimes creakiness. The great diversity of realizations of the glottalized tone when one looks into fine phonetic detail is a matter of great interest from an experimental point of view.

The left graph of Figure 54 shows level F_0 curves of Tone 5 in Experiment 2. This feature differs from qualitative results (Experiment 1) where all three speakers had a rise in their F_0 curves. It can be surmised that the reason why Tone 5 in this experience is flatter might have to do with the difference of overall F_0 . Mean F_0 is higher (for Speaker M1) in Experiment 2 than in Experiment 1, by about 10 Hz (7%). Hence in Experiment 2, its ‘high’ target is already reached early on in the syllable, and so it does not need to be rising, unlike in Experiment 1. Be that as it may, the most important feature of Tone 5 is still preserved. That is the position at the top of F_0 range in comparison with all remaining tones.

O_q of Tone 5 is spread out over a large area from 30% to 80%. Three areas can be distinguished within this large scatter zone: (i) a densely populated

area in the middle, ranging from 42% to 60%; (ii) a much more sparsely populated area from 70% to 80%; (iii) and finally, two isolated items in lowest position from 30 to 40% (as shown in the graph on the right-hand side of Figure 54). Seen in this light, O_q can be said to be located essentially in the middle of range.

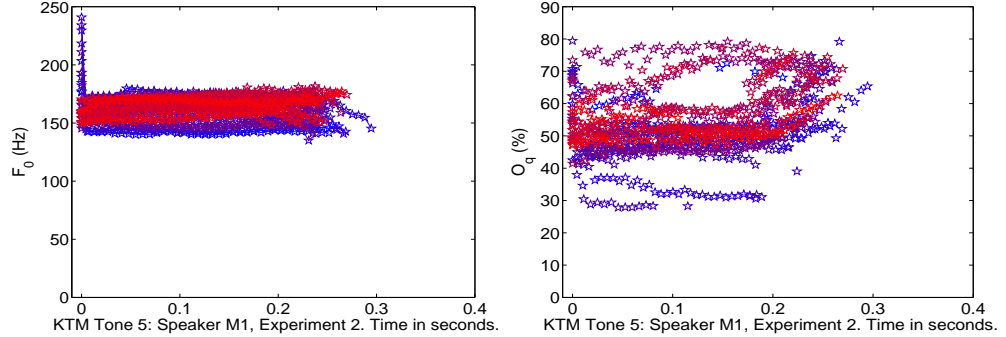


Figure 54 – The curves of F_0 and O_q for Tone 5, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right (Individual items: one curve per item).

Tones 6 and 7, the two tones found in stopped syllables, differ from tones in smooth syllables in term of duration. The same scale is used for the stopped tones in Figures 55 and 56 as for the smooth tones in Figures 49, 51, 52, 54. This brings out visually the great difference in duration between these two tonal subsystems. The stopped tones are about half as short as smooth tones. Their duration is on the order of 0.1 to 0.15 seconds.

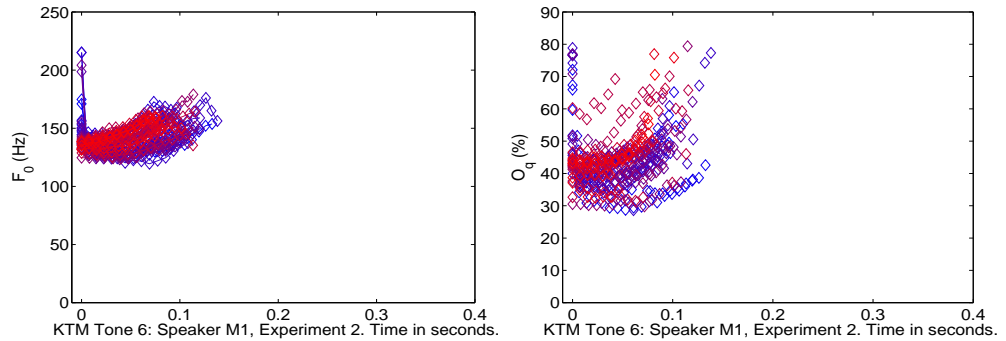


Figure 55 – The curves of F_0 and O_q for Tone 6, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right. (Individual items: one curve per item)

Dealing with Tone 6 first, Figure 55 (the left graph) shows a rising F_0 similar to that for Tone 3. The rise extends from a minimum of 125 Hz to a maximum of 170 Hz, an increase of about 36%. O_q values (in the remaining graph of Figure 55) are spread out from 30 to 80%. Among them, the most values are concentrated in middle and lower-middle areas around 40%. Values over 50% usually occur in the later portion of the rhyme. This could be evidence of soft phonation at the offset of voicing, but given the tendency

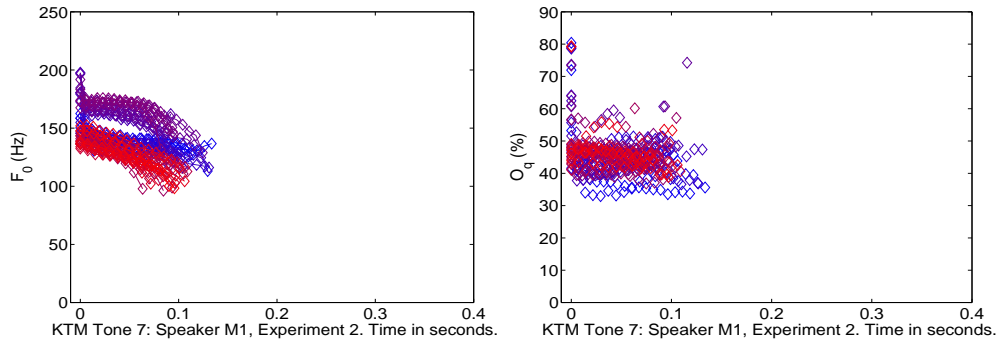


Figure 56 – The curves of F_0 and O_q for Tone 7, in the case of Speaker M1, Experiment 2: (i) F_0 in the left; (ii) O_q in the right. (Individual items: one curve per item)

for F_0 and O_q to correlate positively, the interpretation of the O_q curve needs to take the movement in F_0 into account.

In contrast, *falling* looks like the unmistakable distinguishing property of Tone 7. The left graph of Figure 56 shows individual F_0 curves which can be divided into three sets according to the order of appearance as follows. (i) Towards the beginning of the recording (blue curves), the curves consist of a gradual decrease, from an initial value of about 150 Hz. (ii) Later on in the recording (violet curves), the speaker switches to a somewhat different behaviour, with an onset that is clearly higher than 150 Hz, and a sudden drop towards the end of the rhyme. (iii) Towards the end of the recording (red curves), onset becomes low again, as in (i), but the slope remains strong, as in (ii).

This is a puzzling situation, as if the speaker's behavior shifted in the course of the elicitation session. One hypothesis would be that the entire carrier sentence is realized differently at different points in the session, but a look at the first syllables in the carrier sentence does not reveal an equivalent pattern in three subsets, unlike target items. What could be the reason for this?

Once more, it seems that the influence of Vietnamese is being felt here. Comparing with the representation of Vietnamese tone system by James Kirby [61, p. 386], it is easy to recognize that the first subset of realizations of KTM Tone 7 (those at the beginning of the session, shown as blue curves) looks exactly similar with Vietnamese D2, which, like KTM Tone 7, is one of two tones found in stopped syllables. Hence, we can surmise that there was a confusion between KTM Tone 7 and Vietnamese D2 at the beginning of recording (blue curves). Then the speaker became aware of this, and adjusted to provide syllables that carry KTM Tone 7: those are the violet curves. The curves in red look like the speaker lapsed into Vietnamese pronunciation again towards the end of the session.

Thus, changes in speaker behavior in the course of a recording session do not necessarily follow a regular course – say, decreasing attention and involvement, and accordingly gradually decreasing mean fundamental fre-

quency. There can be rapid shifts in behavior: a sort of phonological code-switching.

O_q values do not show the same type of clear pattern as the F_o curves; statistical processing will be necessary (in future research) to bring out findings in this respect. Pending such exploration, it is tempting to interpret the fact that O_q has a relatively broad range of variation as indicating simply that the phonation type associated to KTM Tone 7 is phonologically unspecified: there is no precise specification as to the phonation type with which this tone is realized; what matters is the contrast in pitch with Tone 6.

3.2.3 *Summary*

After specific descriptions of KTM tone system, reporting on measurements of two parameters (F_o and O_q) in two experiments, I can propose the following conclusions. Firstly, it is confirmed that KTM is a tonal dialect distinguishing 5+2 tones: five contrastive tones in smooth syllables and two contrastive tones in stopped syllables. Despite the fact that the tones of smooth syllables and those of stopped syllables constitute two sub-systems, they are numbered sequentially for convenience, labelling them from 1 to 7.

Secondly, O_q helps to discriminate between Tone 4 (a glottalized tone) and the four other tones (modal tones) by a remarkably low O_q on Tone 4.

Thus, KTM is similar to various languages of mainland Southeast Asia (Burmese, Vietnamese...) in that its tones are distinguished not only by pitch, but by a complex of pitch and phonation types [61], [81], [99]. In outline:

- Tone 1 is a low-mid level tone
- Tone 2 is a falling tone
- Tone 3 is a rising tone
- Tone 4 is a glottalized tone
- Tone 5 is a top-high tone

Duration by itself does not appear to play an important role in contrasts among the five tones of smooth syllables. On the other hand, duration differs greatly between smooth and stopped syllables. The tones in stopped syllables are only about half as long as tones in smooth syllables; this can be ascribed to the influence of obstruent codas. Final /p/, /t/, /c/ and /k/ are phonologically unspecified for voicing (there are no voicing opposition among final consonants in KTM), but these codas are phonetically unvoiced, and appear to exert the full shortening effect that is cross-linguistically associated with final voiceless stops.

Among stopped tones, there appears to be a relatively neat symmetry between a low-rising-checked tone (Tone 6) and a high-falling-checked tone (Tone 7).

The issues encountered in the experimental study need to be recounted in the present Summary. Experiment 2 met unexpected difficulties. The biggest mistake consists in confusion between Tone 1 and Tone 2 in Experiment 2 with two F_o curves that are practically indistinguishable, unlike in Experiment 1 (see Figures 57 and 58 in Chapter 4). For the other tones, the

results of Experiment 2 in term of both fundamental frequency and open quotient look similar to those of Experiment 1, but the contrasts are not as clear as in the experiment with real words (Experiment 1). Especially, the main characteristic of KTM Tone 4 (its glottalization) stands out less clearly in Experiment 2 than in Experiment 1: neither F_0 nor O_q reach as low values for Tone 4 in Experiment 2 as they do in Experiment 1. These issues have to do with the absence of writing system of this dialect, which entails at least two problems. Firstly, the native speakers of this dialect had no initial awareness of their tone system when the study began. This is unlike tonal languages with a writing system, such as Vietnamese. This causes some confusion when speakers are taught to distinguish tones, an unusual task for them, for which the ‘crash-course’ that I provided proved insufficient. Secondly, because of the absence of writing system in this dialect, Experiment 2 used Vietnamese writing to show target syllables. It certainly affected speakers, making them tend to imitate Vietnamese tones when going through this experiment. This weakness needs to be overcome in future work.

DISCUSSION

This chapter consists of four parts. First, it builds on the information provided in the Results section to propose an evaluation of the contribution of various phonetic parameters to the tonal contrasts found in KTM. Second, it dwells on a key issue in Mường studies: the relationships between Mường and Vietnamese, and specifically, how Vietnamese dialectology can provide guidance in Mường studies. A third part of this section is devoted to glottalization – a key topic in the present research –, attempting a general discussion about the linguistic roles of this phonetic universal. Finally, some reflections are set out concerning the social dynamics of KTM.

4.1 GENERAL OUTLOOK OF THE KTM TONE SYSTEM

With a view to offering a convenient overview, Figures 57 and 58 show the averaged trajectories of F_0 and O_q values obtained for speaker M1 in Experiments 1 and 2 respectively.

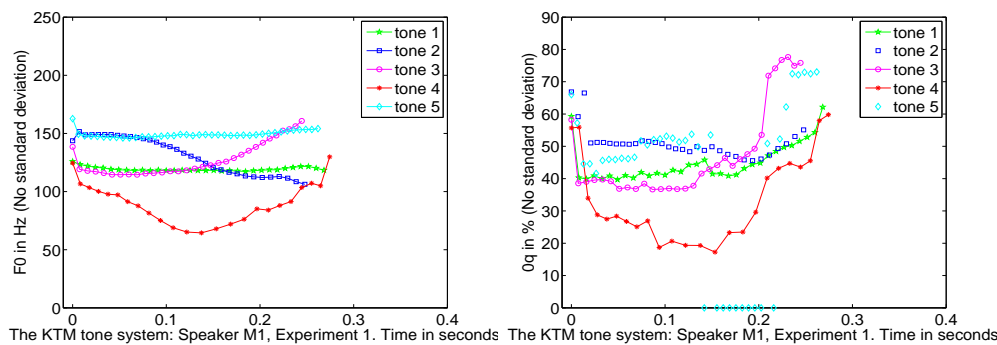


Figure 57 – The averaged curves of fundamental frequency (in the left) and open quotient (in the right) for KTM tone system in the case of Speaker M1, Experiment 1 (no standard deviation), 20 tokens.

Let us review one more time the basic characteristics of KTM tones based on Figures 57 and 58 and the other results in Chapter 3.

Tone 1 is the most level of all five tones, with a moderate final dip corresponding to the offset of voicing. It is usually a little below the speaker's average F_0 and O_q . Tone 2 and Tone 3 are characterized by falling and rising curves respectively; the amplitude of the rise of Tone 3 is not as great as that of the fall of Tone 2. The most remarkable feature of Tone 4 is glottalization, which does not appear at the beginning but in the mid-part of the rhyme, and sometimes there is a lapse into creak. Tone 5 is the highest tone, with F_0 and O_q values always at the top of the speaker's range, so it can be called a 'top-high' tone.

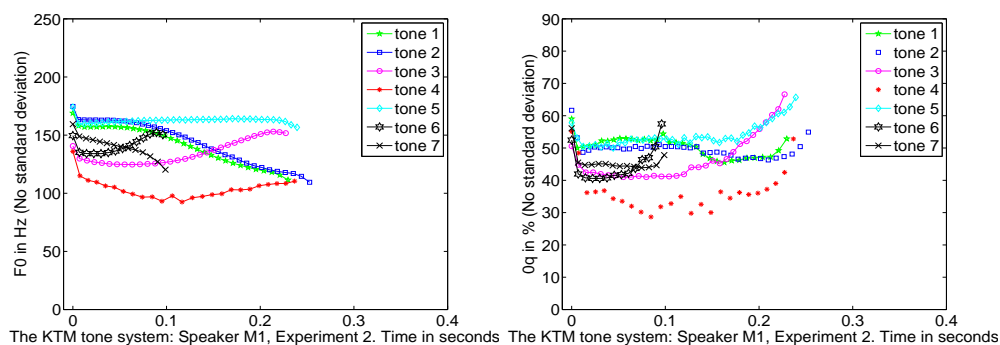


Figure 58 – The averaged curves of fundamental frequency (in the left) and open quotient (in the right) for KTM tone system in the case of Speaker M1, Experiment 2 (no standard deviation), 126 tokens.

Each tonal languages will use different combinations of pitch and phonation type to produce tonal contrasts. An instance of combination between pitch and phonation type to produce tonal contrasts is found in Qingjiang Miao (Ch'ing Chiang Miao), a Black Miao dialect belonging to the Hmong-Mien (Miao-Yao) family. This language has five level tones, among a total of eight tones (the remaining includes two rising and one falling). Tonal space is not sufficient for a good distinction of these five-level tones and duration does not contribute to any contrasts. Phonation cues are used to contribute to a sufficient distance between the tones in the tonal space [66, 67]. In other words, the interaction between pitch and phonation allows for well-dispersed tonal spaces.

In the case of KTM, on first approximation, the evidence examined here suggests that F_0 ensures the bulk of the tonal distinctions, whereas O_q is only relevant insofar as it discriminates the glottalized tone from the remaining four tones (modal tones). Moreover, the glottalized tone is also the lowest of all five tones in terms of pitch, so that pitch and phonation could be said to co-occur; the correspondence between pitch and phonation of this tone system is as expected from a typological point of view: higher pitch corresponds to modal voice and lower pitch correspond to glottalized/creaky voice.

In this light, one could even wonder whether glottalization needs to be granted phonological status at all. The fact that glottalization occurs on the phonetically lowest tone in KTM raises the issue whether this glottalization may not be a low-level phonetic phenomenon associated to a low (or extra-low) phonological target.

But a clear answer in the negative can be provided. The data set out in Chapter 3 reveals that the phasing of glottalization in KTM Tone 4 is precise – much more precise than would be expected if this were simply a low-level phonetic phenomenon. Moreover, the association of glottalization to KTM Tone 4 is highly consistent: glottalization is a robust feature of KTM Tone 4, also found in contexts where speakers raise their voice. Qualitative data observation suggests that the phonetic realization of glottalization is slightly different when the utterance's overall pitch is higher, with less creaky voice and more constriction, but glottalization is still present.

Furthermore, if the glottalized tone were underlyingly (phonologically) specified simply as having an extra-low pitch target, and if glottalization were only a low-level phonetic consequence of this target, then one would expect the lowest tone to be sometimes modal, sometimes creaky, *and also sometimes breathy*. Breathiness is opposite to creaky voice along the continuum of vocal-fold adduction degree; but lapse into breathy voice could be seen as a low-level phonetic consequence of a low pitch target, in the same way as creaky voice.¹

Thus, glottalization phenomena in KTM are consistent with the typological tendency whereby creak is associated with lower pitch, but this typological tendency does not have the power to ‘predict’ glottalization phenomena in KTM in full phonetic detail. The phasing of glottalization inside rhymes carrying Tone 4 does not result from low-level phonetic phenomena. The association of creaky voice to KTM Tone 4 can safely be interpreted as due to a *phonological* specification.

Said differently, pitch plays the main role in the KTM tone system, but it is not the only phonologically relevant parameter. Moreover, further investigation might reveal that Tone 4 is not the only tone in the KTM system that has phonation-type specifications. Glottalization is a conspicuous phonation-type characteristic associated to Tone 4, distinguishing it from the other tones. It may be the case that further research will uncover less obvious phonation-type characteristics in this tone system: characteristics that are less salient (and perhaps less consistent) than the glottalization associated with Tone 4, but which may nonetheless be important to the system.

4.2 VIETNAMESE DIALECTOLOGY AS A SOURCE OF INSPIRATION FOR THE STUDY OF MŨƠNG IN GENERAL AND KTM IN PARTICULAR

KTM is closely related with Vietnamese from a diachronic point of view; moreover, Mường dialects have undergone wave after wave of influence from Vietnamese, strengthening similarities between these language varieties. Vietnamese dialectology, which is overall much more advanced than the study of Mường dialects, can serve as a source of inspiration for the study of Mường in general, and of KTM in particular.

For instance, it appears interesting to see to what extent glottalization in KTM Tone 4 resembles glottalization phenomena in Vietnamese dialects (and in other languages of the Vietic group). In Vietnamese, tone B₂ (*nặng*) and C₂ (*ngã*) are glottalized tones. In a comparison between KTM Tone 4 (based on the result of this study and Figures 52, 57, 58) and Vietnamese B₂ and C₂ (based on the results and figures of [61], reproduced here as Figure

1. This argument could also be extended to Beijing Mandarin. It has often been reported that the Beijing Mandarin third tone is occasionally accompanied by creaky voice. Some authors interpret this as a low-level phonetic consequence of a low pitch target for the third tone. But if this tone’s phonation type were phonologically unspecified, one would expect various types of nonmodal phonation at low pitch in Mandarin, including cases of breathy phonation. It seems that observed cases are tilted towards creak, suggesting that appealing to nonphonological phonetic factors does not tell the full story.

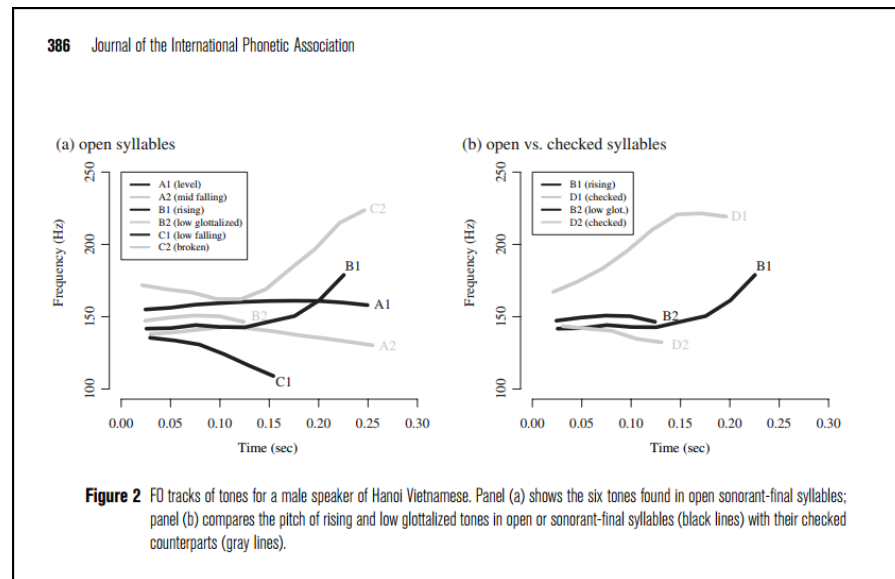


Figure 59 – Figure of F_0 curves of Hanoi Vietnamese tones, from [61, p. 386].

59),² it can be confirmed that KTM Tone 4 differs from the glottalized tones of Vietnamese by at least the following characteristics:

(i) The position of glottalization:

- KTM Tone 4 has *medial* glottalization: the glottalization of KTM Tone 4 happens in the middle part of the rhyme. The tone starts in the lowest part of the speaker's F_0 range: its onset F_0 is the lowest of the five tones; F_0 then decreases sharply, revealing early glottalization; finally, glottalization is relaxed towards the end of the syllable, returning to F_0 values suggestive of phonation mechanism I (chest voice).
- Vietnamese B2 has final glottalization: whereas F_0 curves of KTM Tone 4 can be analyzed into three portions (initial modal portion; glottalized portion; and final modal portion), Vietnamese B2 can be analyzed as consisting of two portions, with glottalization strongest at the *end* of the rhyme.
- Vietnamese C2, similar to KTM Tone 4, has medial glottalization, but the onset and offset F_0 values of tone C2 are not at the same level, whereas those of KTM Tone 4 are close to each other. The figure from [61] reproduced here as Figure 59 shows that the beginning of Vietnamese C2 is high, and its end clearly higher, at the top of the speaker's range.

(ii) The values of F_0 :

Differences in F_0 are relative, and are affected by many factors, such as differences across speakers and experiments. In order to get a feel for this difference, the evaluation of F_0 values of glottalized tones need to be conducted against the background of a comparison with the other tones in the system.

2. I also consulted [93] and [15] to have a good understanding of Vietnamese B2 and C2.

- KTM Tone 4 reaches very low values: it is consistently at the bottom of the speaker's range. Sometimes it can go down to values as low as 40 Hz in the case of speaker M6 (as can be seen in Figure 42).
- Vietnamese B2 as seen on Figure 59 has an F_0 curve in the middle of the speaker's range, lower than A1 and C2, and higher than A2 and C1. The F_0 values at the beginning of the rhyme are thus clearly higher than those of KTM Tone 4.
- The difference between Vietnamese C2 and KTM Tone 4 is even greater: C2 is in the top of the speaker's range at both endpoints of its course, in contrast with KTM Tone 4 which is close to the bottom of the speaker's range throughout. The highest F_0 point of C2 reaches up to levels never seen for KTM Tone 4 (at least in the type of contexts examined here: careful, hyperarticulated realizations).

(iii) Duration

- In KTM, as explained above, pitch and phonation type are both important cues to distinguish between the tones, whereas duration by itself does not appear to play any major role to distinguish between the five tones of smooth rhymes. The duration of KTM Tone 4 (the glottalized tone) is similar to that of the four others; under the experimental conditions used here, the average duration is on the order of 25 centiseconds.
- Vietnamese B2 is *short*. James Kirby's figure of the Vietnamese tones shows that the duration of B2 is only about half that of the other tones. The author comments that "[a]lthough duration has not been shown to be a salient perceptual cue to Vietnamese tone, syllables bearing tones B2 and C1 are often shorter than syllables bearing other tones due to the effects of final glottalization" [61, p. 386].
- The duration of Vietnamese C2 is similar to that of KTM Tone 4.

Overall, although the phonation types of Vietnamese B2 and C2 and of KTM Tone 4 can all be described as glottalized, there are multiple differences between KTM Tone 4, on the one hand, and Vietnamese B2 and C2 on the other. In terms of the tones' position inside the tone system as a whole, the viewpoint defended here for KTM Tone 4 – that glottalization sets it apart from the other four tones of smooth syllables – could not apply to Vietnamese. Vietnamese B2 and C2 appear to have very different ties with the rest of the system. (This point need not be labored here, in view of the wealth of publications about the tone system of the Hanoi dialect of Vietnamese.)

The existence of deep differences does not mean that Vietnamese studies cannot provide inspiration for Mường studies and vice versa. The point that Northern Vietnamese tones "are realized by a complex of pitch and voice quality features" has been made eloquently by specialists of Vietnamese [61, p. 386]; this also applies to Mường, saving specialists of Mường the trouble of laboring a point that has already been made for the closest relative of the Mường language. No less importantly, the issue of which specific phonation type (if any) is actually associated to each tone remains a topical issue in Vietnamese studies; this situation points to the difficulty of reaching

definitive conclusions about which properties of a tone can be considered to be phonological, and how the tone should ultimately be modeled.

[G]lottalization plays an important role in the production and perception of the broken (C₂) and glottalized (B₂) tones. The falling tones (A₂, C₁) have been described by some researchers as accompanied by a breathy voice quality (Thompson 1965; Phạm 2001, 2003); the low falling tone (C₁) has also been described as accompanied by light final laryngealization (Nguyễn Văn Lợi & Edmondson 1998; Michaud 2004; Kirby 2010). [61, p. 386]

These discussions about Vietnamese stimulate the search for other phonation types beside glottalization in KTM.

4.3 GLOTTALIZATION: THE LINGUISTIC ROLES OF A PHONETIC UNIVERSAL

4.3.1 *Introduction*

Phonemic use of glottalization (glottal constriction or creaky voice) is relatively uncommon cross-linguistically. Ladefoged and Maddieson [74, p. 48] mention phonation types as part of their discussion of stop consonants; the examples that they mention of languages making use of creaky voice in association to stops are Hausa and Mazatec. Another example is Hayu [87]. All three are likely to appear as highly “exotic” languages in the eyes of many phoneticians.

In the field of suprasegmentals, on the other hand, glottalization is far from uncommon. The importance of glottalization in prosodic systems of Southeast Asia is emphasized by Brunelle and Kirby [17]; but in Europe too, glottalization in a dialect can correspond regularly to tone in another.

Remarkably, the relation between tone contours in the north [of the Scottish Gaelic area] and glottalization in the south has an exact parallel in Scandinavian as well. Glottalization in Danish (the so-called *stød*) corresponds to word tone in Norwegian and Swedish in the following way: glottalization corresponds to tone 1, absence of glottalization corresponds to tone 2. [112, p. 143]

On the other hand, glottalization is a *phonetic* universal with various uses across languages. In British English, ‘glottaling’ of consonants (especially /t/) is common [103]. In many languages, glottal stop serves as an empty-onset filler: e.g. in French, empty-onset words that have little phonological material tend to be reinforced (set apart from the preceding words) by an initial glottal stop, as in *je m’appelle Yves*, which tends to be realized as /ʒmapelʔiv/ rather than /ʒmapeliv/, or *je m’appelle Anne*, which tends to be realized as /ʒmapelʔan/ rather than /ʒmapelan/ (this observation was communicated by Michel Launey to Alexis Michaud).

... an American English speaker may have a very creaky voice quality similar to the one employed by speakers of Jalapa Mazatec to distinguish the word /jǎ/ meaning ‘he wears’ from the word /já/ meaning ‘tree’ (Kirk et al. 1993 [63]). As was noted some time ago, one person’s voice disorder might be another person’s phoneme (Ladefoged 1983 [73]). [43, p. 383]

The universal nature of creak (also referred to as “creaky voice”; the two labels are used interchangeably here) is especially clear if one uses the classification of phonation mechanisms by Roubeau et al [105, p. 261]: ‘head voice’ is phonation mechanism II; ‘chest voice’ is mechanism I; and ‘creak’ is mechanism zero. While the uses of these three mechanisms differ across cultures, it does not seem that there is any community of speakers in the world that systematically avoids any use of creaky voice. Thus, from a phonetic point of view, creak can be considered as a universal. From a linguistic point of view, its uses range from the phonemic to the ‘paralinguistic’.

4.3.2 *Definitions*

In this thesis, ‘glottalization’ is used as a cover term for glottal constriction, creaky voice, and all the phenomena that range in-between them. Inside the field of glottalization, a distinction is made between, on the one hand, creak and laryngealization, and on the other hand, glottal constriction and glottal stop – keeping in mind that there can be intermediate cases.

4.3.2.1 *Creak and laryngealization*

In the present work, the equivalence between ‘creaky’ and ‘laryngealized’ is adopted, following widespread usage, as reflected e.g. in D. Crystal’s *Dictionary of Linguistics and Phonetics*, where ‘creak’ is defined as follows:

A term used in the phonetic classification of voice quality, on the basis of articulatory and auditory phonetic criteria. It refers to a vocal effect produced by a very slow vibration of only one end of the vocal folds; also known as vocal fry. ... Creaky sounds are also called ‘laryngealized’. [25, pp. 121-122]

Auditory criteria would indeed seem to be central to the notion of ‘creak’, since the term itself refers to an auditory impression. But paradoxically, it seems that phonetic studies focus mostly on production and acoustics. This is a point that I would be interested to investigate further in future: the perceptual counterpart of different types of creak, and how perceptual data can help us find out more about which aspects of the acoustic signal are most relevant to the language users’ perception.

Now coming to more technical phonetic description, Patricia Keating, Marc Garellek and Jody Kreiman propose on the basis of a review of the literature that prototypical creaky voice has three key properties:

- (1) low rate of vocal fold vibration (F_0), (2) irregular F_0 , and
- (3) constricted glottis: a small peak glottal opening, long closed phase, and low glottal airflow. [59]

In detail, the articulatory and acoustic part of the story is not tidy, and departure from canonical realizations are numerous. The list of articulatory and acoustic correlates of creak put together in a recent review is impressive:

[[76] mentioned] low subglottal pressure and glottal flow, slack, thick, compressed vocal folds with a short vibrating length, ventricular contact with the folds, weak or damped pulses, low F_0 , irregular F_0 , period-doubled vibration. Later descriptions (e.g. [20, 42, 64]) added such properties as irregular amplitude, low Open Quotient, skewed glottal pulses, narrow formant bandwidths and sharp harmonics, abrupt closure of the folds, and low spectral tilt. [59]

Given this situation, it may appear necessary to phoneticians to confront the diversity headlong, and propose a phonetic categorization, together with recommendations on tools to tell apart the different categories. Building on earlier three- or four-way classifications of creaky voice into different types, Keating et al. propose a classification of the considerable diversity of phonetic realizations of creaky voice into five kinds. Not all of these five types have the characteristics of prototypical creak. This is a phonetic classification intended to help researchers choose acoustic measurement tools that are appropriate for the kind of creak that they aim to capture; this is not a classification based on the linguistic functions of creak.

- **Vocal fry:** “Creak that is vocal fry with a regular F_0 could instead show higher HNR together with lower formant bandwidths.”
- **Multiply pulsed voice:** “Creak that is multiply pulsed can lack a clear F_0 but instead show subharmonics (resulting in higher values of SHR).”
- **Aperiodic voice:** Creak that lacks the prototypical property of low F_0 ; instead, its properties are low H_1 - H_2 and irregular F_0 .
- **Nonconstricted creak:** “Creak can instead show higher H_1 - H_2 , but still with a low and irregular F_0 .”
- **Tense/pressed voice:** “Creak that is more like tense or pressed voice can have a mid or high, and regular, F_0 .” [59]

The acoustic correlates of these five types are summarized in table form in Figure 60.

To me, this table looks impressive, mentioning measurements that I have never attempted so far, such as measurements of the relative amplitude of different harmonics in the signal. I take it as a word to the wise, that the categories may be in part determined by the tools used in data collection and analysis (in the case of the present study: electroglottography and spectrograms). Use of different exploratory techniques tends to channel the phonetician’s interpretation in different ways. One perspective for future work will consist in examining the distribution of phonetic creak in KTM Tone 4 in terms of the five-way classification of creak proposed by Keating et al.: to

Property >	low F0	irreg F0	glottal constr	damped pulses	sub-harms
Main correlate > Type √	low F0	high noise	low H1-H2	low noise; narrow BWs	high SHR
proto-typical	√	√	√		
vocal fry	√		√	√	
multiply pulsed		√	√		√
aperiodic	NO	√	√		
nonconstricted	√	√	NO		
tense	NO		√		

Figure 60 – Properties characterizing different kinds of creak. A tick means that the property characterizes this type; NO means that it does not; blank means variable or unknown. Reproduced from [59].

what extent one type of creak or other is more common (i) for certain speakers, (ii) in certain speaking styles, (iii) in association with specific segments, and so on. This work could also be extended to occasional cases of glottalization that may be observed for the other tones (phonologically nonglottalized tones) in continuous speech. It could be that the five-way classification proposed by Keating et al. uncovers consistent patterns whereby creak in KTM Tone 4 is modulated, so to speak, by the speakers.

4.3.2.2 *Glottal constriction and glottal stop*

An overview of definitions is proposed by John Esling, placing the term ‘glottal stop’ in historical perspective:

In the phonetic literature, there is a long but not always consistent history of the definition of glottal stops. Holder (1669, pp. 60, 72 [56]) defined a glottal stop as “a stop made by closing the larynx”. Bell (1867, pp. 46, 60 [7]) defined a glottal stop as a glottal catch made with the glottis closed and a catch of the breath as in a cough, while specifying that the linguistic effect of a glottal stop is softer than in a cough. Sweet (1877, pp. 6–7 [109]) also called a glottal stop a glottal catch and defined it as a sudden opening or closing of the glottis. He also claimed that the most familiar example of a glottal catch is in an ordinary cough. Noël-Armfield (1931, p. 107 [97]), Heffner (1950, p. 125 [53]), and Jones (1956, p. 19 [58]) defined a glottal stop

as a closure and opening of the glottis. Jones implied that the glottis must be tightly closed. Ladefoged (1975, p. 46 [72], 1988, p. 50 [75]) states that a glottal stop is made by holding the vocal cords tightly together and also suggests that glottal stops occur in coughs. Laver (1994, pp. 187–188, 206 [77]) defines a glottal stop as a maintained complete glottal closure.

In most modern phonetic literature, therefore, a glottal stop is defined simply as a closed glottis or tightly closed glottis without any reference to the ventricular folds, arytenoid cartilages, or supraglottal cavity activities. [30, pp. 385–386]

The two terms *glottal stop* and *glottal constriction* are considered to refer to two different phonetic phenomena by Sprigg [106, p. 5], but according to Esling these belong to a continuum: “...glottal stops have been observed to occur on a continuum from a weakly constricted glottal stop to a strongly constricted glottal stop (...)” [30, p. 386].

A typical configuration observed in the production of a glottal stop “includes an adduction of the arytenoid cartilages, a complete adduction of the vocal folds, a partial adduction of the ventricular folds, and moderate narrowing of the laryngeal vestibule through its epilaryngeal sphincter mechanism (Esling, 1996, pp. 72–73 [32], 1999, pp. 358–369 [29]; Harris, 1999, 2001 [46, 47])” [30, p. 386].

Phonetically, glottal stops and glottal constrictions occur in many languages, including English [22], French [83], German [65], and Hanoi Vietnamese [90]. In KTM, canonical realizations of the glottalized tone (Tone 4) show laryngealization (lapse into creaky voice) rather than a glottal stop or glottal constriction. Investigations into different kinds of materials (such as narratives, or elicited materials in which the syllables carrying Tone 4 are not under focus) are likely to uncover a wider range of realizations, however. Since Tone 4 is the only glottalized tone in the KTM system, its glottalization can be expected to show a larger field of allophonic variation than in tone systems that contrast two or more glottalized tones (such as Hanoi Vietnamese, for instance).

4.3.3 *Glottalization in KTM*

As has been mentioned at several places in this thesis, glottalization makes Tone 4 well-distinguished from the other four tones of smooth syllables. Glottalization sets apart Tone 4 by two acoustic properties. Firstly, fundamental frequency: Tone 4 is very different from the other tones, as its F_0 values usually drop from under 100 Hz to minimum values below 50 Hz whereas the four remaining tones are consistently inside a range between 100 and 150 Hz. Even without information of open quotient, Tone 4 can be identified as being the lowest tone, at the bottom of speaker’s range, contrasting with the four remaining tones (which are distributed in the area from the middle to the top of speaker’s range). Secondly, open quotient also

tells apart Tone 4 from the others, with values of Tone 4 reaching below 30% whereas all the other tones have O_q values above 40%.

To what extent these two acoustic properties – fundamental frequency and open quotient – contribute to the listeners’ perceptual impression of the tones is a matter for future studies of speech perception to investigate.

As an example of hypotheses to be tested in future work, it seems that creak with complex-repetitive patterns (to which Keating et al. refer as ‘multiply pulsed’) occurs frequently in canonical or hyperarticulated realizations of KTM Tone 4. It is a safe guess that reduction to less canonical creak is frequent in casual speech – a hypothesis which will later be tested through examination of other materials, such as narratives.

4.4 SOME NOTES ABOUT THE SOCIAL DYNAMICS OF KTM

KTM is a non-written language variety inside a dialect continuum characterized by in-depth influence from the national standard – Hanoi Vietnamese. The Mường language remains locally vigorous; at the same time, Hanoi Vietnamese is clearly a prestige language, whose acquisition is key to success in school – opening prospects for a brighter future and better work opportunities outside Kim Thượng.

As mentioned in Chapter 1, KTM is defined as the dialect of the (few) villages of Kim Thượng commune. In terms of social status, this is the strongest group in this area with a population that makes up more than 75% of the total. The rest is divided among four groups: Dao, H’ông, Tày, Kinh, which differ sharply from one another. Locally, the nation’s “majority” ethnic group, the Kinh, are thus among the smallest minorities in terms of population. The Dao and Mường each form a community with festivals and habits of its own; members of other groups, such as Kinh, Tày and H’ông people, are new settlers or married into one of the two local communities. Its peripheral location in geographic terms would seem to make KTM more likely to endure than other varieties of Mường. Contact between the Mường and Dao is not strong: Dao group lives separately in areas that have more difficult terrain. Local people refer to the respective habitats of the Mường and Dao as “khu trên” and “khu dưới” (meaning upper and lower zone, respectively). This can be seen as the acknowledgment of a difference between the two groups’ current settlement: the lowlands to the Mường, the highlands to the Dao. It could be that this description also reflects that the two should remain distinct: affirming their distinctness. Due to their current living habits and standards, the Dao people are not engaged in particularly active trade exchanges, nor are they greatly integrated to the network of health and education services. Therefore, despite geographic proximity, contact between the Dao and Mường in Kim Thượng is relatively loose. As to ethnic H’ông and Tày, they live interspersed within the Mường community but their number is very small and studying their interaction with their surroundings, and their potential influence on KTM, would require specific case studies (a monograph about a household, for instance) which I can

hardly hope to conduct, for want of having any knowledge of their mother language.

As for ‘Kinh’ newcomers to the area, their influence is linked to their status as representatives of the ethnic majority of Vietnam. From my field observation, local people in Kim Thượng have a positive attitude to their mother language. All the local people that I have met use their mother language for day-to-day communication. Individuals who belong to a different ethnic group but have married into the KTM community also tend to fall in line with local people by learning and using this language, yielding to a demand that seems overwhelming. Therefore, KTM dialect keeps the role of the strongest language in this community to this day. On the other hand, it seems clear that KTM is influenced by the “Kinh” language (Hanoi Vietnamese) much more than by any other. There is strong contact in many aspects. Firstly, the development of commercial relations created an opportunity for Mường people in Kim Thượng to use Vietnamese to communicate and trade. Vietnamese now becomes the main language in the market, standardly used between people unacquainted with each other. Secondly, the role of Vietnamese is becoming increasingly important in the KTM community, as the language of education, health services, and administrative affairs. Nowadays, children are educated in Vietnamese from the beginning of their schooling. Vietnamese is also the language used for television, radio, and other media mass. Because KTM in particular and Mường language in general have no writing, they have to use Vietnamese for administrative documents to address the affairs of their community. Last but not least, local people recently tend to go out of their area, mainly going into “Kinh” communities, to seek employment when farm work ends. Those reasons explain why almost 100% Mường people in Kim Thượng are able to use Vietnamese (to different extents). The stage is thus set for strong contact between KTM dialect and Vietnamese, leading to language variation and change in KTM.

The predicament encountered when running Experiment 2 is a case in point: as mentioned in the report about this experiment, the consultants tend to identify KTM tones with their phonetically closest equivalent in Standard Vietnamese, resulting in confusion of Tone 1 and Tone 2. The consultants’ great difficulties when required to steer clear of Standard Vietnamese tones and focus solely on the pronunciation of their mother tongue shows that they have much greater metalinguistic awareness about their second language than about their mother tongue. For them, metalinguistic notions of ‘tone’ are so strongly linked to Vietnamese (the language which they learned to analyze into phonemes in the process of learning Vietnamese writing) that special training is required for them to develop an awareness of their mother tongue’s categories. This could seem paradoxical, but is probably widespread in communities whose mother tongue is an unwritten language in a context of unequal bilingualism.

Although sociolinguistics and language contact were not intended as the main topic of the present research, it is clear that these are simply unavoidable. Any study of KTM needs to include special precautions and verifi-

cations concerning possible interferences from Vietnamese. Since there is no way to deny or avoid this complexity, the way forward seems to be to embrace the complexity and give it a central place in the research. As a speaker of Vietnamese, and a ‘Kinh’ newcomer to the community (no matter how much effort is spent making my presence inconspicuous and unobtrusive), I am fortunate to be a witness to cases of code-switching, blending and other phenomena, which I intend to document through field notes, with the possible mid- and long-term prospect of a sociolinguistic survey using the methods gradually developed in sociolinguistics from [113–115] to [68–70].

CONCLUSION

This dissertation is based on an experimental study aiming to provide a better understanding of phonetic dimensions that are known to be important cues to linguistic tone: fundamental frequency, which has pitch as its perceptual counterpart, and glottal open quotient, a parameter that relates to phonation types. A special focus of this thesis consisted in taking up the challenge of studying phonation of glottalized tones.

As a result of fieldworks, the documentation of KTM tone system can be said to reach a stage where fundamental tonemic issues can be considered solved, and new phonological and phonetic questions pertaining to tonation can be explored. Specifically:

(i) A plentiful and diversified corpus. It contains vocabulary, narratives, conversations, and songs, in addition to tone experiments. Beyond the results reported in the present volume, this set of data holds potential for further studies on both intrinsic and extrinsic issues of tones (such as co-articulation, intonation, influences of speaker attitudes on tones). It is hoped that it can also become useful resources for other research fields, e.g. lexicology and grammar.

(ii) A experimental study applied `PEAKDET`, a MatLab script available from the COVAREP GitHub repository [27], as a basis for modelling phonetic characteristics of tones. This contributes to setting a new trend for the use of electroglottography as part of the experimental setup in phonetic studies of tone in Vietic and in other languages of Vietnam.

(iii) The qualitative and quantitative results reveal phonetic characteristics of KTM tone system, in terms of two basic parameters of fundamental frequency and glottal open quotient.

On the other hand, this study also has at least four limitations. (i) Lack of symmetry between the two experiments because of the difficulty in seeking minimal sets. A larger amount of vocabulary will help for this. (ii) Shortage of female data because the more noisy electroglottographic signal from women is a big challenge when using the tools chosen in the present study. (iii) Last but not least, due to my beginner level in Matlab, I have not yet been able to apply (and develop further) tools for the modeling of creaky portions of rhymes. (iv) Statistical data analysis is still sorely lacking.

There remains much work to do, from the short term to the long term.

(i) Overcoming the limitations mentioned above: finding new minimal sets and collecting more data from female speakers to arrive at more balanced coverage and hence better-grounded conclusions; performing further studies on fundamental frequency, and phonation types and their various phonetic correlates; the relationship and interaction between them; last but

not least, supplementing this study by perceptual experiments,¹ with statistical data treatment.

(ii) Extending the investigation to:

- Intrinsic and co-intrinsic effects such as effect of vowels and consonants.
- Intonation effects: variation of glottalization depending on position in sentence, position in the syntactic structure; and pragmatic factors: emphasis played on a word, as well as speaker attitudes,
- Laryngography: images of the larynx during phonation, as a desirable complement to electroglottography, like some pioneering papers [15, 31].

(iii) Towards diachronic study: comparing KTM with other Mường dialects and furthermore with languages in Việt - Mường (Vietic) group, especially Vietnamese.

This dissertation is only a start in the work of describing and analyzing KTM tone system in synchrony. The many issues that remain constitute a captivating research program for the long term.

1. Work on the perception base of tonal contrasts has developed in Southeast Asia recently, largely thanks to technological developments facilitating F_0 and O_q calculation. Among specialists in this field, I am aware of the work of Burnham [19], Svantesson [108], Abramson et al [1–4], Zsiga & Nitisaroj [118], Brunelle et al [10, 11, 13, 16, 18], Kirby [60, 62], and Gruber [44]

Part II

APPENDICES

This second part of the volume groups Appendices containing data and analyses from fieldwork, including: (i) phonemic analysis (vowels and consonants), (ii) word list, (iii) recorded materials, (iv) scripts, and (v) additional figures.

APPENDIX 1: THE VOWELS AND CONSONANTS OF KTM

This thesis essentially deals with tones. In fieldwork, however, working out the target dialect's inventory of vowels and consonants is a preliminary to further work. This Appendix sets out observations about the vowels and consonants of KTM.

6.1 CONSONANTS

6.1.1 *Initial consonants*

An inventory of initial consonants is shown in Table 9. Among the phonemes in Table 9, /w/ is as a newcomer to the system, appearing on loanwords from Vietnamese, which have initial /v/ in Vietnamese.

Table 9 – Initial consonants inventory

	Labial	Dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Plosive	p p ^h	t t ^h		ʈ		k k ^h	ʔ
Nasal	m		n		ɲ	ŋ	
Trill			r				
Fricative	β		s			ɣ	h
Approximant	w				j		
Lateral							
Approximant			l				
Implosive	ɓ		ɗ				

6.1.2 *Final consonants*

Table 10 – Final consonants inventory

p	t	c	k
m	n	ɲ	ŋ
w	j		

This system resemble the inventory of final consonants which proposed by Ferlus for Làng Lữ Thổ [41], a dialect considered as part of “Central

Mường”¹ by Henri Maspero [85]. The two palatal consonants /c/ and /ɲ/ in Làng Lô Thổ, according to Michel Ferlus, “are found (i) in borrowings of Vietnamese words with final /c ɲ/ (orthographic *ch* and *nh*), which originate in velars following high front vowels, and (ii) in a few inherited words, such as *kăc*²⁴ (Viet. *cắt* [kătD1]) ‘to cut’ and *sɲ*²⁴ (Viet. *rắn* [rănB1]) ‘snake’”. The origin of **c** and **ɲ** in Kim Thượng Mường seems to be the same; interestingly, the list of inherited words is longer: see Table 11. In addition, palatals originating in the fronting of velars are not restricted to the context of a following high front vowel (a structure illustrated by *as* /kip1/ ‘Vietnamese’) but also appear after mid-low front vowels, e.g. in /sep1/ ‘green’. These observations are observed with a low degree of certainty, as I am still a beginner in attempting Vietic historical phonology; for want of any real certainty I have not systematically indicated in Table 11 if a word is inherited or borrowed.

Table 11 – Palatal final consonants **c** and **ɲ** in inherited vocabulary and in loan words

Num.	Kim Thượng Mường	Vietnamese	English	Note
1	/mɛɲ1/	nhANH, mau	quick	
2	/tɛɲ1/	đAN	to weave	
3	/nwɛɲ2/	nhỏ, bé	small	
4	/kip2/	gần (nơi)	near to	
5	/kɛɲ4/	cắN	to bite	
6	/rɛɲ4/	cái đố, lò	fish basket	
7	/tʰɛɲ4/	rắN	snake (general)	
8	/nyɲ4/	nắN, vắT	to squeeze	
9	/kac6/	cắT, chắT	to cut	
10	/kʰac6/	sắT	iron	
11	/kac7/	cắT	sand	
12	/ke4 lac7/	dây lắT	bamboo strips	
13	/tɛc7 ɣuəɲ4/	đắT, để xuốNg	put	
14	/mic7/	mắT	honey	
1	/sep1/	xANH	green	
2	/kip1/	KINH	Vietnamese	another word: /taw4/
3	/mɛɲ3/	mắNH	strong	
4	/tɛɲ4/	đắNH (nhau)	to fight	

Table 12 – Vowel inventory

i	ɯ	u
e	ɣ/ɣ̃	o
ɛ	a/ǎ	ɔ
iæ		
iə	ɯə	uə

6.2 VOWELS

Basically, the vowel system of Kim Thượng Mường is close to Vietnamese with nine monophthongs /i e ɛ a ɯ ɣ u o ɔ/, three falling diphthongs /iə ɯə uə/, and two pairs of long and short vowels /ɣ-ɣ̃/ /a-ǎ/. As one goes into subphonemic detail, some interesting differences appear. For instance, Kim Thượng Mường has a falling diphthong [iæ], in complementary distribution with [ɛ]. The former appears before all final consonants other than the two palatals, /ɲ/ and /c/. Examples include: /tɛɲ4/ ‘to fight’, /tɛc/ ‘to put’, as opposed to /riæp7/ ‘sandals’, /kʰiæt7/ ‘lightning’, /kiæk7/ ‘armpit’, /kiæm1/ ‘ice-cream’, /tiæn4/ ‘narrow’, and /tiæŋ2/ ‘nail’.

1. Henri Maspero used the term “Mường” to refer to all languages of the Vietic 2 group except “Annamese”(Vietnamese)

APPENDIX 2: WORD LIST

The following is based on the word list of Michel Ferlus, an expanded version of the EFEO-CNRS-SOAS word list for linguistic fieldwork in Southeast Asia ¹. This word list was used throughout fieldwork for starting approaches of KTM tone system and for collecting KTM vocabulary as well, but it was not the main task of this study, so only a third of the word list has been collected. It is provided here with IPA transcription. The word list given here is exported from a database, sorting the list by tone, then by consonant.

UID	English	Vietnamese	(KTM)
1442	to comb one's hair	chải tóc	caj1 t ^h āk6
1924	two	hai	han1
1856	monday	thứ hai	t ^h u4 han1
72	cave	hang, động	haŋ1
2033	exit	ra	(ti1) ha1
203	flower	bông, hoa	poŋ1, hwa1
2287	to clap one's hands	vỗ tay	hwan1 ho1
1058	family	gia đình	ja1 diŋ5
201	thorn	gai	kaj1
512	orange	cam	paj4 kam1
1271	span	gang tay	kaŋ1 t ^h āj1
133	cast iron, pig iron	gang	kaŋ1
463	eggplant, aubergine, brinjal	cà tím	ka1 tim4
473	tomato	cà chua	ka1 cuə1
1873	next year	năm sau, năm tới	k ^h aŋ1 năm1
15	star	ngôi sao	ŋoj1 k ^h aw1
1440	comb	lược	kiəŋ1 k ^h aŋ1
554	cockscorn, crest	mào (gà)	maw1 ka2
51	south	(phía) nam	p ^h iə4 nam1
264	areca palm tree	cây cau	kǎn1 naŋ1
510	Annona squamosa	na	paj4 na1
366	wet rice field (in the plains)	ruộng nước	na1 rak7
1011	mother's elder brother's wife	mợ, bác	na1

1. The EFEO-CNRS-SOAS word list is available online: https://halshs.archives-ouvertes.fr/halshs-01068533/file/EFEO_CNRS_Ferlus_SEAWordList_v1.pdf

UID	English	Vietnamese	(KTM)
1857	tuesday	thứ ba	t ^h u ₄ pa ₁
1927	three	ba	pa ₁
1363	to dilute	hòa, pha	p ^h a ₁
1694	knife	dao con	jaw ₁ , raw ₁
2409	to wash (rice)	vo gạo	taj ₁ kaw ₄
2429	to carry on a pole	gánh	tam ₁
164	coal	than	t ^h an ₁
1124	ear	tai	t ^h aj ₁
1882	tomorrow	ngày mai	ten ₄ βaj ₁
2204	to vomit	nôn mửa, ói mửa	βa ₁
495	jute	cây đay	k ^h n ₁ đăj ₁
1962	a hundred	một trăm	mut ₇ kă ₁ m ₁
883	python (large)	trăn	kă ₁ n ₁
1320	to be drunk	say rượu	k ^h ăj ₁ raw ₃
1726	go hunting	đi săn	ti ₁ k ^h ăn ₁
1889	after	sau	k ^h ăw ₁
12	moon	mặt trăng	mă ₆ t ₆ păn ₁
1849	waxing moon	trăng tròn	păn ₁ kwôn ₅
1868	year	năm	ră ₁ m ₁
180	root	rễ cây	răj ₁ k ^h yn ₁
1859	thursday	thứ năm	t ^h u ₄ răn ₁
1929	five	năm	ră ₁ m ₁
299	to husk rice	xay lúa	săj ₁ lô ₃
348	to grind	xay (thành bột)	săj ₁
2364	to pick one's teeth (clean)	xia răng	tăm ₁ t ^h ăn ₁
2669	right	(bên) phải	pen ₁ tăm ₁
2315	(physical pain)	đau (thể xác)	tăw ₁
1162	arm, upper arm	tay	t ^h ăj ₁
1135	tooth	răng	t ^h ăn ₁
1172	hand	bàn tay	pan ₅ t ^h ăj ₁
2690	who? (in who is that man?)	ai đây ?	?ăj ₁ re ₃
1319	to chew betel	ăn trầu	?ăn ₁ pu ₅
1286	to eat meat	ăn thịt	?ăn ₁ puk ₆
2398	to chew betel	ăn trầu	?ăn ₁ pu ₅
391	edible bamboo shoots	măng	βăn ₁
940	children (sons and daughters)	con cái	kôn ₁ kaj ₄
2674	inside	bên trong	k ^h a ₄ kôn ₁

UID	English	Vietnamese	(KTM)
976	daughter	con gái	kɔn1 kaj4
1193	buttocks	mông, đít	kʰɔŋ1
948	young (people)	trẻ	mon3 nɔn1
2692	whence? (in where has he come from?)	ở đâu đến ?	ʔɿ1 nɔ1 ten4
2693	where? (in where are you going?)	đi đâu?	ti1 nɔ1 re3
2694	when? (in when are you going?)	khi nào anh đi?	puə1 nɔ1 ʔuə3 ti1
2275	to kill	giết, sát hại	pɔ1
939	children (general)	trẻ con, con nít	kɔn1 rwɔj1
2179	to be sated, replete	no	rɔ1
516	rose-apple	roi, mận	rwɔj1
1470	shuttle	con thoi	kɔn1 tʰwɔj1
828	bee	ong mật	ʔɔŋ1
1029	father's father	ông nội	ʔɔŋ1 noj3
1034	husband's father	bố chồng	ʔɔŋ1 ja3
2672	above	bên trên	kʰa4 ken1
2069	to forget	quên	kwen1
1545	head-cushion, pillow	cái gối	ken1 kon4
2195	to be well, in good health	khoẻ mạnh	kʰwe1 meŋ3
2641	quick	nhanh, mau	meŋ1
2751	blue (indigo)	xanh chàm	seŋ1
1467	to weave	dệt, đan	teŋ1
2577	fragrant	thơm	hɿm1
1869	last year	năm trước, năm ngoái	nă1 hɿ1
297	cooked rice	cơm	kɿm1
409	cooked rice	cơm	kɿm1
2596	smooth	trơn nhẵn	kɿn5
1282	armful	một ôm	mot7 ʔɿm1
546	paws, feet (of animals)	bàn chân	pan5 cɿn1
1100	(human) limbs	tay chân, tứ chi	cɿn1 tʰăj1
1203	leg	cái giò, cái chân	ke4 cɿn1
1205	foot	bàn chân	pan5 cɿn1
179	tree	cây cối	kɿn1 kon4
2703	<negation> (spoken)	không	kʰɿŋ1
49	west	(phía) tây	pʰiə4 tɿj1

UID	English	Vietnamese	(KTM)
2319	to make a hole in	đâm lỗ	tʃm1 lo5
303	to pound rice	giã gạo	tʃm1 kaw4
108	marsh	đầm lầy	tʃm1 rʃm3
1097	body	thân thể	tʰʃn1 tʰe5
2572	hot (of peppers)	cay	tʰʃn1
539	bird	chim (trống/mái)	cim1
733	bird	chim	cim1
2691	what? (in what are you eating?)	ăn gì ?	ʔăn1 ci1 re3
1074	Vietnamese	(người, dân tộc) Kinh, Việt	mon3 kij1
2687	this (house)	(cái nhà) này	ke4 na5 ni1
1253	heard	tim	tim1
1981	to go	đi	ti1
492	lemon	chanh	ciæŋ1
2547	jealous	ganh, ghen tị	ɣiæn1 ti3
382	bamboo (general term)	tre, pheo, mai, giang	kʃn1 kiæ1
2747	black	đen	siæŋ1
2670	left	bên trái	pen1 ciəw1
2688	that (house)	(cái nhà) đó, kia	kiə1
856	water leech	con đĩa, tắc	tiə1
2221	swollen	sưng	kʰuŋ1
255	cucumber	dưa chuột, dưa leo	juə1 re5
481	watermelon	dưa hấu	juə1 hʃw4
1461	mirror	cái gương, kính	ke4 kuəŋ1
30	dew	sương, sương móc	kʰuəŋ1
1811	to saw	cưa	kʰuə1
20	rain	mưa	muə1
1951	twenty	hai mươi	han1 muən1
2560	fresh (vegetables)	(rau) tươi	tʰak6 tʰuəj1
48	east	(phía) đông	pʰiə4 doŋ1
94	river	sông	kʰoŋ1
1215	hair (of body)	lông (mình)	loŋ1
565	feather	lông	loŋ1
1154	lip	môi	moj1
163	lime	vôi	pon1
2646	shallow	nông cạn	roŋ1

UID	English	Vietnamese	(KTM)
1429	pair	đôi, cặp	mot7 toj1
798	prawn	tôm	t ^h om1
1059	descendant	con cháu, người nối đôi	t ^h on1
974	son's son	cháu trai nội	t ^h on1 tuə4
934	man (vir)	đàn ông	păj1 ʔon1
1026	mother's father	ông ngoại	ʔon1 ɲwaj3
319	to plant a tree	trồng cây	βon1 kʰn1
1339	boil	(nước) sôi	rak7 huj1
973	daughter in law, son's wife	con dâu	kɔn1 ju1
963	wife (legitimate wife)	vợ, vợ hợp pháp	ju1
992	younger brother's wife	em dâu	ʔun4 ju1
1666	bow	cái cung	kuŋ1
854	earthworm	con trùn, giun	kɔn1 k ^h un1
2645	deep	sâu	k ^h u1
347	flour	bột	puŋ1
2616	wet	ẩm ướt	ʔɿm5 t ^h un1
2548	happy	bằng lòng, hài lòng	βuj1
2874	bell	chuông	cuəŋ1
793	crab	cua	kuə1
1845	noon	giữa trưa, đúng Ngọ	kuə1
801	eel	lươn	luən1
2484	buy	mua	muə1
2101	to go down, descend	xuống	suəj1
1219	bone	xương	suəŋ1
548	tail	đuôi	tuəj1
2883	you	mi, mày, bạn, cậu	ja2
277	to sow rice (into hole)	tra thóc	ka2 lɔ3
538	chicken	gà trống/mái	ka2 k ^h on4
727	chicken, domestic fowl (gen- eral)	gà nhà	ka2 ɲa5
2452	to stir	quậy, khuấy	kwaj2
2722	smaller than	lớn thua, nhỏ hơn	nwaŋ2 hɯn1
2032	enter	vào	(ti1) paw2
1032	father's mother	bà nội	pa2 noj3
1030	mother's mother	bà ngoại	pa2 ɲwaj3
1462	cotton (material)	vải bông	paj2 poŋ1

UID	English	Vietnamese	(KTM)
2140	to answer, reply	trả lời, đáp	pa2 lyj5
2625	clean	sạch sẽ	kwaŋ2
2679	far, distant (from)	xa (nơi)	sa2
96	cascade	thác	taj2
1218	skin	da	ta2
428	banyan (figus trees)	cây đa	kʰn1 ta2
612	stag, deer	hươu	βaŋ2
769	to hunt	săn mồi	kʰǎn1 moj5
2667	behind	đàng, đằng sau	kʰa4 kʰǎw1
1374	hash	băm, bằm, vằm	pǎm2
765	to fly	bay	pǎn2
62	plain	đồng bằng	toŋ2 pǎŋ2
2599	thick	dày	tǎj2
1505	sew	may khâu	βǎn2
2876	I, me	tao, tôi, tớ, mình	hɔ2
219	grass	cỏ	kɔ2
537	tiger	hổ, cọp (đực/cái)	mɔŋ2
536	elephant	voi (đực/cái)	pwɔj2
1739	to kill an animal	giết thú	pɔ2
892	tortoise	rùa	rɔ2
1495	to thread	xâu kim	sɔ2 kim1
2750	red	đỏ	tɔ2
1434	necklace	vòng cổ	wɔŋ2 kɔk7
478	spiny amaranth	rau dền gai	tǎk6 cen2
2098	to go up, ascend	lên	leŋ2
2183	to suck	bú	mem2
584	goat	dê	te2
2619	small	nhỏ, bé	nweŋ2
2441	to hang	treo	tɛw2
2031	to climb	leo, trèo	tɛw2
485	ginger	gừng	kyŋ2
2025	to swim	bơi	pyj2
2118	to feel	sờ, sờ thấy	ry2
469	long gourd	bầu dài	βǎw2
2640	slow	chậm	cǎj2
2507	rich	giàu có	cǎw2

UID	English	Vietnamese	(KTM)
2425	to hold in the hand	cầm trong tay	kʰm2 kɔŋ1 tʰǎj1
2296	to push	xô, đẩy	tʰj2
2678	near to	gần (nơi)	kiŋ2
2600	thin	mỏng	hiæn2
190	branch	cành, nhánh	kiæn2, ɲiæn5
1618	oil lamp	đèn dầu	tiæn2
1708	chopping-log or block	thót	ke4 tiæn2
1815	nail	đinh	tiæn2
2511	divide	chia	ciə2
2870	gong	chiêng	ciəŋ2
2449	to carry in one's arms	bế, bồng, bưng	puŋ2
2410	to wash (hands)	rửa tay	tʰuə2 tʰǎj1
421	tubers (general)	củ	ko2
719	dragon	rồng	roŋ2
1519	vegetable garden	vườn, vườn rau	won2 tǎk6
533	buffalo	trâu (đực/cái)	ku2
579	buffalo, water buffalo	trâu	ku2
2537	dumb	câm	lun2
1543	mosquito net	cái mùng, cái màn, màn chống muỗi	tuj2
2534	one-eyed	một mắt, chột	tun2
158	cinders, ashes	tro	βun2
2567	acid	chua	cuə2
1376	soup	xúp, cháo	caw3
2081	stupid	ngu, đần	jaŋ3
2163	to cry, weep	khóc	jam3
2451	turn back, turn around	quay lại	pɣ2 laŋ3
1998	return	trở lại	pɣ2 laŋ3
1194	knee	đầu gối	ko4 laŋ3
991	elder brother's wife	chị dâu	maŋ3 ju1
286	rice seedlings	cây mạ	kʰn1 ma3
993	elder sister's husband	anh rể	maŋ3 cʰw4
985	brother (woman speaking)	anh em	ʔun4 maŋ3
2419	to grab	nắm, bắt	nam3, pǎt6
1507	to patch up clothes	vá quần áo	pa3
1053	companion/friend	bạn bè	pan3 pɛ5

UID	English	Vietnamese	(KTM)
1370	alcohol	rượu	raw3
1722	ball, projectile, bullet	đạn	tan3
2791	ghost	hồn	waj3
983	elder sister	chị	?aj3
988	sister (man speaking)	chị em	?aj3 ?un4
2036	to run	chạy	căn3
2643	high	cao	jăn3
58	high	cao, trên cao	jăn3
1627	cradle	cái nôi	ke4 kăn3
2744	white	trắng	kăn3
877	dragon fly	con chuồn chuồn	kun1 kwăj3
18	set (the sun sets)	lặn (mặt trời)	măt7 pxj5 lă3
2647	heavy	nặng nề	năn3 ne5
2565	bitter	đắng	tăn3
2591	curved	cong	wăn3
1134	throat	cổ họng	ko1 hơj3
2119	to speak	nói	kơ3
2538	stutter	cà lăm, nói lắp	kơ3 ơơj3
1548	mattress	đệm, nệm	tem3
2654	strong	mạnh	mej3
1516	market	chợ	cơ3
356	crock	bình, chậu	cỡw3
66	slope	dốc	ky3, cơj4
2366	to dig	bới, bươi đất	pxj3 tởt6
2050	wait	đợi, chờ	txj3
2213	smallpox	đậu mùa	dỡw3 muə5
2049	to get up	đứng dậy, dậy	jởn3
2064	to wake up	thức dậy	jởn3
414	maize	bắp, ngô	k ^h ỡw3
1839	evening	chiều (trước 6-7 giờ), chiều tối, hoàng hôn	k ^h wỡj3
1051	marriage (woman speaking)	lấy chồng	lỡj3 lỡw3
2259	to marry (of a man)	lấy vợ	lỡj3 ju1
957	husband	chồng	lỡw3
2260	to marry (of a woman)	lấy chồng	lỡj3 lỡw3
178	bushy, dense	rậm rạp	rỡm3

UID	English	Vietnamese	(KTM)
505	Phaseolus radiatus, mung bean	đậu xanh	tʃw3 siɛŋ1
1252	kidney	thận	tʰɿn3
1010	mother's elder brother	cậu, bác	ʔɿw3
995	father's elder brother	bác trai	ci3 kʰa5
1481	to spin (on a spindle)	se sợi	kiæw3 paj5
2677	beside	ở cạnh, ở bên cạnh	pen1 kiæŋ3
2297	to pull	kéo vào	kiæw3 paw2
2630	cold	lạnh, rét	liæŋ3
2289	to nip, pinch	béo, véo	miæw3
969	mother	mẹ	miæ3
1038	wife's mother	mẹ vợ	miæ3
1131	mouth	miệng	mom5, miæŋ3
1036	husband's mother	mẹ chồng	miæ3
2648	light (in weight)	nhẹ nhàng	ŋiæn3
378	well	giếng	ciæŋ3
710	horse	ngựa	ŋwə3
982	elder brother	anh (trai)	ʔwə3 tuə4
1832	play	chơi	ti1 jon3
183	trunk (of trees)	thân cây	ko3 kʰn1
2869	drum	trống	koŋ3
2515	steal	ăn trộm, đánh cắp	ʔăn1 lom3
2037	flee	trốn	lon3
933	man (homo)	người ta, con người	kaj3 mon3
867	termite	con mối	mon3
1191	belly (abdomen)	bụng	pon3
2651	large	rộng rãi	ron3
1546	blanket	mền, chăn	toj3
2428	to carry on the head	đội trên đầu	toj3 ken1 kok6
227	inter-node	đốt, lóng	tot6
1001	father's younger brother	chú	cu3
125	stone	đá	kʰu3
2405	dust	bụi	pun3
1244	stomach	dạ dày	puŋ3 kaj4
1463	silk (material)	lụa	luə3
298	rice soup	cháo	caw4
1146	forehead	trán	caŋ4

UID	English	Vietnamese	(KTM)
285	to harvest rice	gặt lúa	haj4 lɔ3
2470	to wish	muốn, ước	haw4
531	female (animal)	cái, mái	kaj4
304	bran	cám	kam4
1157	chin	cầm	kan4
786	fish	cá	ka4
958	widowed (husband)	góa, góa phụ	pa2 kwa4
294	husked rice	gạo xay, gạo thô, gạo nâu	kaw4
987	younger sister	em gái	ʔun4 kaj4
1866	month, or moon	tháng, tháng âm	kʰan4
76	way, path, unpaved road	đường	kʰa4 kaj4
1678	arrow	cái tên	tɔn1 lan4
198	leaf	lá	la4
388	slim bamboo	nứa	law4
1144	cheek	má	ma4
1277	height gauge (to measure somebody's height)	sải tay	paj1 tʰaj1
2485	sell	bán	pan4
999	father's elder brother's wife	bác gái	pa4
211	fruit	trái, quả	paj4
1343	to grill	nướng	ran4
1354	to fry	chiên, rán	ran4
1355	to fry	rán	ran4
2284	to stamp one's foot	đá, đập	ta4
1236	urine	nước tiểu	rak7 taj4
1077	Tai	người Thái	mɔn3 tʰaj4
2618	dry	khô	kʰo1, tʰaw4
1932	eight	tám	tʰam4
308	to winnow rice	sàng gạo	tʰa4 kaw4
1841	morning	sáng	βan4
2568	salty	mặn	căm4
763	to run	chạy	căn4
2460	to truss up	trói (ai)	kwǎj4
1860	friday	thứ sáu	tʰu4 kʰǎw4
1930	six	sáu	kʰǎw4
472	manioc, tapioca, cassava	sắn	kʰăn4

UID	English	Vietnamese	(KTM)
625	gnaw	gặm, nhấm	k ^h ăm ₄
1228	blood	máu	măw ₄
2288	(finger) flick	búng tay	năj ₄ t ^h ăj ₁
2107	to hear	nghe thấy	ηăj ₄ ʔan ₄
2072	to think (about)	nghĩ (tới)	ηăm ₄
2066	to snore	ngáy	ηăn ₄
1861	saturday	thứ bảy	t ^h u ₄ păj ₄
1931	seven	bảy	păj ₄
2623	short	ngắn, vắn	păn ₄
2644	low	thấp	păn ₄
59	low	thấp, dưới thấp	păn ₄
2030	bathe	tắm	t ^h ăm ₄
2176	to break wind	địt, đánh rắm	tɛɲ ₄ t ^h ăm ₄
1394	dress	áo dài, váy, đầm	wăn ₄
1327	fish sauce	nước mắm	rak ₆ βăm ₄
2849	divination	thuật chiêm bói, bói	βwɔj ₄
535	dog	chó (đực/cái)	cɔ ₄
594	dog	chó	cɔ ₄
93	stream	suối	hɔn ₄
155	smoke	khói	k ^h wɔj ₄
2605	difficult	khó khăn	k ^h ɔ ₄
1185	(finger) nail	móng tay	mɔɲ ₄ t ^h ăj ₁
1341	hot	nóng	nɔɲ ₄
2629	hot	nóng	nɔɲ ₄
1177	fingers	ngón tay	ηɔn ₄ t ^h ăj ₁
2103	to look at	nhìn	ηɔ ₄
32	wind	gió	sɔ ₄
2404	broom	cái chổi	tɔm ₄ t ^h ɔk ₆
2178	to be hungry	đói	tɔn ₄
1308	to drink alcohol	uống rượu	ʔɔɲ ₄ raw ₃
1303	to drink water	uống nước	ʔɔɲ ₄ rak ₆
1332	salt	muối	βwɔj ₄
321	mushroom (growing on the ground, in the soil)	nấm (mọc dưới đất)	cen ₄
517	carambola, starfruit	khế	k ^h e ₄
2083	to hope	mong, hi vọng	pen ₄
1996	come	đến, lại	ten ₄

UID	English	Vietnamese	(KTM)
2415	to bite	cắn, gặm	kɛɲ4, kʰăm4
2637	big and fat	to béo	pɛw4
1785	fish basket, bamboo basket, catching cage	cái đó, lồng	rɛɲ4
2276	to fight	đánh trả, chống trả	tɛɲ4 pa1
878	snake (general)	rắn	tʰɛɲ4
2607	sparse, scanty	thưa, thưa thớt	hɤn4
1466	paper	giấy	jɤj4
2602	easy	dễ dàng	lɤj4
2656	new	mới (mẻ)	mɤj4
354	to squeeze	nặn, vắt	nɤɲ4
977	son in law, daughter's husband	con rể	kɔn1 cɤw4
994	younger sister's husband	em rể	ʔun4 cɤw4
1326	vinegar	dấm	jɤm4
280	to transfer out rice seedlings	cấy lúa	kɤn4 lɔ3
777	egg	trứng	kɤɲ4
40	thunder	sấm	kʰɤm4
2799	prayer	câu kinh, lời khấn	kʰɤn4
2270	to bury	chôn	pɤm4
2614	bad	xấu, dở	sɤw4
2639	ugly	xấu	sɤw4
1362	to soak	ngâm	tɤm4
1342	warm	ấm	ʔɤm4
1933	nine	chín	cin4
2557	cooked	(thịt) chín	cin4
815	head louse	chí, chấy	ci4
816	body louse	rận	ci4
464	wax gourd, stone pumpkin	bí đao	pin4 đaw1
1572	latrine	cầu tiêu, nhà xí	ho4 si4
1004	father's younger brother's wife	thím	ʔi4
563	wing	cánh	kiæɲ4
1389	undress	cởi (quần áo)	kiæ4 quən5 ʔaw4
352	cake	bánh	piæɲ4
2716	similar	giống, tương tự	piæn4 ra5

UID	English	Vietnamese	(KTM)
296	non-glutinous rice	gạo tẻ	kaw ₄ riæ ₄ , kaw ₄ căm ₁
2653	narrow	hẹp	tiæn ₄
941	adopted child	con nuôi	kɔn ₁ ciəm ₄
826	ant	kiến đen	kiæn ₄
226	sugar cane	mía	miæ ₄
866	cricket	con dế	tiæn ₄
2655	weak	yếu ớt	ʔiəw ₄ ʔɪt ₇
514	pineapple	dứa, thơm	paj ₄ jwə ₄
972	son	con trai	kɔn ₁ tuə ₄
984	younger brother	em trai	ʔun ₄ tuə ₄
2884	it, he	nó, hắn, anh ấy, chị ấy	ʔuə ₃
863	grasshopper	châu chấu	co ₄
2563	rotten	(thịt) thối, (trái) bủn, (quả) ủng	ɲuk ₆ hon ₄
340	mortar (for pounding rice)	cối giã gạo	kon ₄ tɿm ₁ kaw ₄
2411	wash (hair)	gội (đầu)	kon ₄ kok ₆
1192	back	lưng	noŋ ₄
1178	thumb	ngón cái	ŋon ₄ kaj ₄
1928	four	bốn	pon ₄
1858	wednesday	thứ tư	tʰu ₄ pon ₄
332	to cook rice (water)	nấu cơm	ro ₄ kɿm ₁
1335	to cook (general)	nấu nướng, nấu ăn	ro ₄ ʔăn ₁
1336	to boil (water)	đun, nấu nước	ro ₄ rak ₆
1383	to cook (in soup / sauce)	nấu canh	ro ₄ kiɛŋ ₁
1266	navel	cái rún, rốn	tʰon ₄
1840	evening	tối (sau 7 giờ)	tʰon ₄
2196	to be ill, sick	đau ốm	ʔom ₄ tăw ₁
2401	to blow	thổi	hun ₄
658	wild boar, wild pig	heo, lợn rừng	kun ₄
598	pig (domestic)	heo, lợn	kun ₄
752	owl (general)	cú mèo	ku ₄ mew ₅
305	to winnow paddy	sảy lúa	kum ₄ lɔ ₃
1717	gun, rifle	súng	kʰuŋ ₄
231	banana	chuối	cuə ₄
792	tadpoles	con nòng nọc	kuən ₄ kɔn ₁

UID	English	Vietnamese	(KTM)
843	butterfly	bướm, bướm bướm	puəm4
1800	hammer	búa	puə4
2521	saliva	nước miếng, nước nhãi	rak7 jaj5
2622	long	dài	jaj5
2621	big	lớn, to	ka5
2634	fat	to	ka5
2721	bigger than (in his house is bigger than mine)	lớn hơn	ka5 hɤn1
471	radish	củ cải	ko1 kaj5
276	to broadcast rice	gieo vãi lúa	kwaj5
947	old man	người già	mon3 k ^h a5
952	old (of people)	già	k ^h a5
1356	to stir-fry; to sauté	xào	kɤn1 k ^h aw5
327	to cultivate a rice field	làm ruộng	la5 na5
1780	casting net	lưới, chài	ke4 laj5
1753	ploughshare	lưỡi cày	la5 j5 kǎn1
240	sweet potato	khoai lang	ku2 law5
1828	work	làm việc	la5 wek7
1510	village (under one headman)	thôn, làng, xóm	laŋ5
1575	ladder	cái thang	ke4 man5
1528	house	nhà	ɲa5
2043	fall (oneself)	ngã té	ɲa5
2675	outside	bên ngoài	pen2 ɲwaj5
33	typhoon	bão	paw5
1585	table	cái bàn	ke4 pan5
1589	(meal) tray	cái mâm	ke4 pan5
1806	plane	cái bào	ke4 paw5
935	woman	đàn bà	pǎj1 pa5
430	mango	xoài	swaj5
129	gold	vàng	waŋ5
1426	shoes or clogs, footwear	giày, dép	jǎj5, riɛp7
1428	to wear shoes	mang giày, đi giày	ti1 jǎj5
1751	plough	cái cày	ke4 kǎn5
341	hand holding pestle	chày tay	k ^h ǎj5
2062	to sleep	ngủ	tǎj5
2580	straight	thẳng	t ^h ǎŋ5
847	larva	tằm, nhộng	t ^h ăm5

UID	English	Vietnamese	(KTM)
2481	to seek	tìm kiếm	tiz cɔ̌5
2187	to stand	đứng	cɔ̌5
746	heron	con cò	kɔn1 kɔ̌5
2584	round	tròn	kwɔn5
2632	thin	ốm gầy	kɔm5
122	river gravel, fluvial gravel	sỏi, sạn	kʰwɔ̌j5
1250	liver	gan	lɔm5
823	mosquito	muỗi	mwɔ̌j5
2044	drop (object)	đánh rơi	rwɔ̌j5
1830	idle	rảnh, rỗi	ʔɿ5 rɔ̌5
2649	hard	cứng	rɔ̌5
721	bat	dơi	sɔ̌5 we5
486	garlic	tỏi	kɔ1 twɔ̌j5
1438	bracelet	vòng tay, cườm	wɔ̌j5 tʰăj1
47	rainbow	cầu vồng	kʰwɔ̌j5 wɔ̌j5
561	beak	mỏ	βɔ̌5
1188	breast (female)	vú	mem5
2650	soft	mềm	mem5
622	rat	chuột	re5
1838	night	đêm	tem5
587	cat	mèo	mɛw5
1650	raft	bè	pɛ5
2092	to draw	vẽ	we5
1863	period	buổi, giờ	jɿ5
2673	below	bên dưới	kʰa4 jɿ5
4	sun	mặt trời	măt7 pɿj5
1	sky	trời	pɿj5
1246	breathe	thở	tʰɿ5
2626	dirty	nhơ, bẩn, nhớp	bʰn5
1380	to stew	hầm	hʰm5
1328	oil (edible)	dầu	jʰw5
19	cloud	mây	mʰn5
1431	ring (for finger)	nhẫn	ɲʰn5
738	goose	ngỗng	ŋʰŋ5
1401	wrap a turban	quấn, chít khăn	kwʰn5 kʰăn1
2661	full	đầy	tʰj5
513	papaya	đu đủ	paj4 tʰw5 jʰw5

UID	English	Vietnamese	(KTM)
967	father	cha	ʔɤj5
1035	wife's father	cha vợ, bố vợ	ʔɤj5
135	lead	chì	ci5
1483	(spun) thread	chỉ, sợi	kʰɤj1 ci5
2412	rub and wipe	kì cọ	ki5 cɔ3
1705	axe	rìu	ke4 riw5
490	tea plant	cây chè	kʰn1 ciæ5
192	branch (small)	nhánh nhỏ	ɲiæɲ5 nweɲ2
1763	shovel	cái xẻng	siæɲ5
2255	to bear	sinh, đẻ	siɲ1 tiæ5
809	to lay eggs	đẻ trứng	tiæ5 kuɲ4
953	to give birth	sinh đẻ	siɲ1 tiæ5
556	to bring forth	đẻ, sinh	siɲ1 tiæ5
2711	few (in there are few houses)	ít	ʔiæɲ5
2226	diarrhoea	ỉa chảy	kʰɔj1 ʔiæ5
115	sea	biển	biæɲ5
518	galangal, galanga	riềng	ko2 riɲ5
1596	spoon	muỗng, thìa	siə5
1591	dish	cái đĩa	tiə5
2489	money	tiền bạc	tʰiæɲ5 pak7
551	hornes (of buffalo)	sừng	kʰuɲ5
2109	to smell	ngửi (thấy)	ɲuɲj5
165	forest	rừng	ruɲ5
2707	some	vài, nhiều	tu5
2250	to be pregnant	chửa, có mang, có bầu, có thai	ko4 cuə5
1549	bed	giường, sàng	cuəɲ5
2125	to abuse	chửi rủa, lăng mạ, lăng nhục	cuəj5
1323	cane sugar	đường mía	ɗuəɲ5 miə4
256	coconut palm	cây dừa	kʰn1 juə5
1554	doorway	cửa	kuə5
2431	to carry together	khênh, khiêng	kʰuəɲ5
2676	middle, between	ở giữa	ʔɤ1 kʰuə5
1934	ten	muối	muəɲ5
1764	harrow	cái bừa	puə5
494	grape-fruit, pomelo	bưởi	paj4 puəj5

UID	English	Vietnamese	(KTM)
2305	to fear, to be afraid	sợ	ruəj5
107	lake	hồ	ho5
355	pot (for cooking rice)	nồi nấu cơm	noŋ5 ro4 kɤm1
1245	lung	phổi	pʰoj5
2102	to see	thấy	ro5
63	mountain, hill	núi, đồi	ton5
131	copper	đồng	toŋ5
1478	reel	guồng quay tơ	won5
432	guava	ổi	paj4 ʔoj5
567	nest	ổ, tổ chim	ʔo5 cim1
868	termitary	tổ mối	ʔo5 mon3
1791	fishnet	rổ, vó	cum5
2657	old	cũ	ku5
141	fire	lửa	kuj5
2544	dwarf, short	lùn	lun5
1402	rain-hat, coolie-hat	nón, mũ	mu5
1127	nose	mũi	muj5
441	corypha palm	cây cọ	kɤn1 mun5
2399	betel quid	miếng trầu	pu5
268	betel	trầu	pu5
128	mud	bùn	pun5
2265	to hug	ôm	ʔu5
2663	half empty	lưng, vơi	nuə5
1540	room	căn buồng, phòng	puəŋ5
2551	sad	buồn rầu	puən5
811	fly	ruồi	ruəj5
1599	chopsticks (eating)	đũa	tuə5
2316	to cut	cắt, chặt	kac6
132	iron	sắt	kʰac6
1149	tongue	lưỡi	laj5
2749	yellow	vàng	wan5
50	north	(phía) bắc	pʰiə4 ɓăk6
2408	to wash (clothes)	giặt quần áo	căt6
2073	to understand	hiểu, thấu	măt6
855	land leech	con vắt	păt6
462	vegetable (general)	rau	tăk6
1105	hair of head	tóc	tʰăk6

UID	English	Vietnamese	(KTM)
2168	to call	gọi (ai)	hək6
784	to sing (of animals and birds)	(chim) hát	hwət6
291	chaff, straw	rơm	tʰək6
223	straw, rice straw	rơm	tʰək6
212	seed (of fruit)	hạt, hột	ʔək6
1879	yesterday	hôm qua	ha2 jɤp6
117	dry land, mainland	đất liền	tʰt6 liən5
60	earth, land	đất	tʰt6
2369	epilate	nhổ lông	pik6 loŋ1
2414	wring	vắt khô	ciəp6
1888	before	trước	kwək6
2665	in front	đàng, đằng trước	kʰa4 kwək6
1890	moment	chốc, lát, một lúc	ma2 lok6
2611	good	tốt, hay	tʰot6
1831	help	giúp, hộ, đỡ	la2 jup6, la2 cuə1
2556	raw	(thịt) sống	juk6 kʰoŋ4
1231	flesh	thịt	juk6
1334	meat	thịt	juk6
474	luffa, sponge gourd	muốp	puəp6
484	snake gourd	muốp rắn	puəp6
2175	to belch	ợ	ʔuək6
2569	sharp (to the taste, like pineapple)	chát sít	cat7
1744	rope	dây thừng, chèo, chạc	cak7 kun5
119	sand	cát	kac7
1444	fan	cái quạt	ke4 kwat7
2717	different	khác nhau	kʰak7 ra5
396	bamboo strips (long, narrow strips of bamboo)	dây lạt	ke4 lac7
254	groundnut, peanut	lạc, đậu phộng	lak7
130	silver	bạc	pak7
1592	bowl (small size)	cái bát, chén, đọi	pat7
239	beans (general)	đậu, đỗ	pap7 tɔ5
2165	tear	nước mắt	rak7 măt6
83	water	nước	rak7
2525	snot	nước mũi	rak7 muj5

UID	English	Vietnamese	(KTM)
1098	corpse	xác	sak7
2047	throw (a stone)	ném, quăng	tap7
1166	shoulder	vai	wak7
2426	to carry on the shoulder	vác trên vai	wak7
680	macaca	khỉ	wak7
743	crow	quạ	kɔn1 ʔak7
1214	heel	gót chân	cɿn1 kwǎk7/ kwǎk7 cɿn1
1114	eye	mắt	mǎt7
1388	put on, wear	mặc (quần áo)	mǎk7
2317	to whittle, cut out, to shape	gọt, chạm gọt	kɔt7 pɔ5
1151	neck	cổ	kɔk7
17	rise (the sun rises)	mọc (mặt trời)	mǎt7 pəj5 mɔk7
2570	sweet	ngọt	ŋwɔt7
2334	to peel	bóc, lột vỏ	pɔk7
1243	intestines	ruột	rwɔt7
2090	to read	đọc	tɔk7
1622	hearth	lò, bếp lửa	pep7 kuj5
1571	kitchen	bếp	pep7
295	glutinous rice, sticky rice, sweet rice, waxy rice	gạo nếp	kaw4 rep7
901	frog	ếch, ngóe	ʔek7
2450	put	đặt, để xuống	tɛc7 yuəŋ4
42	lightning	chớp	cɿp7
1256	gall	mật gan	mɿt7
841	honey	mật	mic7
739	wild duck	vịt trời	wit7
1167	armpit	nách	kiæk7
45	the lightning strikes	sét đánh	kʰiæt7 tɛp4
1430	sandals	dép	riæp7
2536	deaf	điếc	tiæk7
134	tin	thiếc, kẽm	tʰiæk7
2418	hold	giữ	cuk7
1168	chest	ngực	ŋuk7
528	male (animal)	đực, trống	tuɔk7, kʰoŋ4
82	place	nơi, chỗ	tuwək7

UID	English	Vietnamese	(KTM)
1280	handful	một nhúm, một nắm	mut7 cot7
1101	head	đầu, trốc	kok7
906	snail	ốc	?ok7
1281	two handfuls	vốc	βok7
199	bud	chồi, mầm	mut7
1922	one	một	mut7
921	squirrel	sóc	cuət7
317	to harvest by hand	tuốt lúa (bằng tay)	tuət7 lɔ3

APPENDIX 3: RECORDED MATERIALS

This Appendix provides information about all the audio files collected during field trips in the course of this study. It includes 49 audio files (the total recording time is around eight hours and a half) with 4 categories: experiment for tone, vocabulary, narrative, and dialogue. My dissertation has only used a small part of this database. The rest is intended for use in further work for my Ph.D. research. There are plans for hosting this data in the Pangloss Collection http://lacito.vjf.cnrs.fr/pangloss/index_en.htm, which already hosts Mường materials collected by Michel Ferlus from 1983 to 1996.

The information here includes:

- (i) Name of audio files. The format is WAV. The identifier for each document begins with <crdo-> (the former name of the archive) followed by the ISO code of the language in the Ethnologue catalogue of languages (the code for Mường is MTQ), then a short identifier of the document in capital letters (the underscore is allowed as a separator).
- (ii) Duration of audio files.
- (iii) Code of consultants who made those audio files: “F” is Female and “M” is Male; the numbers from 1 to 6 refers to the order in which consultants participated in the investigation. Detailed information about consultants was mentioned in Chapter 2.

n°	name of WAV file	duration	consultant code
1	crdo-MTQ_KTM_F1_TONEXP_V1	00:07:57	F1
2	crdo-MTQ_KTM_F1_TONEXP_V2	00:08:47	F1
3	crdo-MTQ_KTM_F2_TONEXP_V1	00:07:12	F2
4	crdo-MTQ_KTM_F2_TONEXP_V2	00:07:12	F2
5	crdo-MTQ_KTM_F3_TONEXP1	00:07:12	F3
6	crdo-MTQ_KTM_F3_TONEXP2	00:06:37	F3
7	crdo-MTQ_KTM_F4_TONEXP1	00:09:10	F4
8	crdo-MTQ_KTM_F4_TONEXP2	00:05:47	F4
9	crdo-MTQ_KTM_F5_TONEXP1	00:22:03	F5
10	crdo-MTQ_KTM_F5_TONEXP2	00:07:30	F5
11	crdo-MTQ_KTM_M1_TONEXP_V1	00:09:20	M1
12	crdo-MTQ_KTM_M1_TONEXP_V2	00:07:45	M1
13	crdo-MTQ_KTM_M2_TONEXP1	00:16:46	M2
14	crdo-MTQ_KTM_M2_TONEXP2	00:10:18	M2

n°	name of WAV file	duration	consultant code
15	crdo-MTQ_KTM_M3_TONEXP1	00:06:57	M3
16	crdo-MTQ_KTM_M3_TONEXP2	00:05:25	M3
17	crdo-MTQ_KTM_M4_TONEXP1	00:09:41	M4
18	crdo-MTQ_KTM_M4_TONEXP2	00:04:36	M4
19	crdo-MTQ_KTM_M5_TONEXP1	00:07:46	M5
20	crdo-MTQ_KTM_M5_TONEXP2	00:09:26	M5
21	crdo-MTQ_KTM_F1_VOC1	00:08:44	F1
22	crdo-MTQ_KTM_F1_VOC2	00:19:43	F1
23	crdo-MTQ_KTM_F2_VOC	00:19:16	F2
24	crdo-MTQ_KTM_F3_VOC	00:19:32	F3
25	crdo-MTQ_KTM_F4_VOC	00:25:34	F4
26	crdo-MTQ_KTM_F5_VOC	00:20:12	F5
27	crdo-MTQ_KTM_M1_VOC	00:16:32	M1
28	crdo-MTQ_KTM_M2_VOC	00:24:37	M2
29	crdo-MTQ_KTM_M3_VOC	00:19:25	M3
30	crdo-MTQ_KTM_M4_VOC1	00:05:53	M4
31	crdo-MTQ_KTM_M4_VOC2	00:11:50	M4
32	crdo-MTQ_KTM_M5_VOC	00:17:26	M5
33	crdo-MTQ_KTM_F5_Narrative1	00:04:09	F5
34	crdo-MTQ_KTM_F5_Narrative2	00:03:24	F5
35	crdo-MTQ_KTM_F5_Narrative3	00:02:25	F5
36	crdo-MTQ_KTM_F5_Narrative4	00:02:33	F5
37	crdo-MTQ_KTM_F5_Narrative5	00:02:42	F5
38	crdo-MTQ_KTM_F5_Narrative6	00:03:04	F5
39	crdo-MTQ_KTM_F1M1_DIALOGUE	00:02:12	F1M1
40	crdo-MTQ_KTM_F1M1_DIALOGUE2	00:01:13	F1M1
41	crdo-MTQ_KTM_F1M1_DIALOGUE3	00:01:25	F1M1
42	crdo-MTQ_KTM_F1M1_DIALOGUE4	00:06:29	F1M1
43	crdo-MTQ_KTM_F1M1_DIALOGUE5	00:05:01	F1M1
44	crdo-MTQ_KTM_F1M1_DIALOGUE6	00:16:21	F1M1
45	crdo-MTQ_KTM_M1M6_DIALOGUE	00:06:19	M1M6
46	crdo-MTQ_KTM_M1M6_DIALOGUE2	00:06:29	M1M6
47	crdo-MTQ_KTM_M1_TONEXP2_REPET	00:21:33	M1
48	crdo-MTQ_KTM_M6_TONEXP1	00:03:16	M6
49	crdo-MTQ_KTM_M6_TONEXP2	00:32:12	M6

APPENDIX 4: SCRIPTS

As mentioned in the previous chapters, Appendix 4 presents the scripts used in this study. I hope that these materials will become a useful reference for readers who take interest in this topic and want to engage in experimental study about Mường dialects or other interesting languages.

After processing EGG signal by PEAKDET script as described in chapter 2, the following scripts, written with the help of my supervisor Alexis Michaud, also run in MatLab to sort the data into sets of tone and then plot all their F_0 and O_q values in a matrix. The content of scripts is as follows:

9.1 SCRIPT FOR SORTING THE RESULTS OF PEAKDET

Script:

```
% Script to sort the Muong data of experiment 1 (real-
words).

% clearing workspace
clear

% Indicate the speaker.
spe = 'M1';

% Speaker:
if spe == 'M1'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-MTQ_KTM_M1_TONEXP1_REPET';
elseif spe == 'M5'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-MTQ_KTM_M5_TONEXP1';
elseif spe == 'M6'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-MTQ_KTM_M6_TONEXP1';
end

% Loading the data
load (address)
datat1 = []; datat2 = []; datat3 = []; datat4 = [];
    datat5 = [];

% Counting the number of processed rhymes
sz = size(data);
```

```

% Information about the structure of the data in the
  experiment.
t1 = [01 06];
t2 = [02 07];
t3 = [03 08];
t4 = [04 09];
t5 = [05 10];

% Counters for number of syllables.
n1= 0;   n2 = 0;   n3 = 0; n4 = 0; n5 = 0;

for i = 1:maxnb
    % clearing the variable containing the processed
      item
    clear datafile

    % loading a token
    datafile = data(:, :, i);

    % calculating the number of periods
    period_nb = length(datafile(:, 1));

    % getting the identifier of the rhyme: the ITEM
      number
    item = numb(i);

    %%%% Analyzing the label
    % changing the identifier to a string
    itemstr = num2str(item);

    % isolating last character: number of repetition.
    rep = itemstr(length(itemstr));
    % isolating character before last: position of the
      rhyme inside the carrier sentence
    pos = itemstr(length(itemstr) - 1);
    % removing 2 last characters from the string.
    itemstr = itemstr(1:(length(itemstr) - 2));
    itemnb = str2num(itemstr);
    % test on position inside sentence. 1 for first
      syllable of frame; 3 for target syllable; etc.
    if pos == '3'
        % test on the item number
        if ismember(itemnb, t1)
            n1 = n1 + 1;
            chem1(n1) = item;
            for q = 1:10
                for r = 1:period_nb
                    datat1(r, q, n1) = datafile(r, q);
                end
            end
        elseif ismember(itemnb, t2)
            n2 = n2 + 1;

```

```

chem2(n2) = item;
for q = 1:nbcol
    for r = 1:period_nb
        datat2(r,q,n2) = datafile(r,q);
    end
end
elseif ismember(itemnb,t3)
    n3 = n3 + 1;
    chem3(n3) = item;
    for q = 1:nbcol
        for r = 1:period_nb
            datat3(r,q,n3) = datafile(r,q);
        end
    end
elseif ismember(itemnb,t4)
    n4 = n4 + 1;
    chem4(n4) = item;
    for q = 1:nbcol
        for r = 1:period_nb
            datat4(r,q,n4) = datafile(r,q);
        end
    end
elseif ismember(itemnb,t5)
    n5 = n5 + 1;
    chem5(n5) = item;
    for q = 1:nbcol
        for r = 1:period_nb
            datat5(r,q,n5) = datafile(r,q);
        end
    end
else
    disp('Error: syllable not found in inventory')
end
end
end
end

% save as: ...._SORTED.mat
filename = [address '_SORTED']
save(filename)

```

9.2 SCRIPTS FOR PLOTTING THE DATA

9.2.1 *Tone by tone*

This script helps to plot F_0 and O_q results of each individual tone with output of 4 graphs:

1. F_0 result: raw data with all curves of each individual item.
2. O_q result: raw data with all curves of each individual item.
3. F_0 result: average data with standard deviation.

4. O_q result: average data with standard deviation.

Script:

```
% To plot the F0 and Oq curves of each individual tone
    in a matrix by many ways

% Opening a figure, clearing it, and setting a value
    for font size.
% DO NOT CHANGE THE ORDER OF THESE LINES.

% clearing workspace
clear

% Indicate the speaker.
spe = 'M1';

% Speaker:
if spe == 'M1'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-
        MTQ_KTM_M1_TONEXP1_REPET_SORTED';
elseif spe == 'M5'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-MTQ_KTM_M5_TONEXP1_SORTED';
elseif spe == 'M6'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-MTQ_KTM_M6_TONEXP1_SORTED';
end

% Loading the data
load (address)

figure(1)
%clear
clf
%setting a value for font size
h = axes;
set(h, 'FontSize', 19)

figure(2)
clf
h = axes;
set(h, 'FontSize', 19)

% Choice of data set
data = datat1;
nbtokens = length(nonzeros(data(1,1,:)));
```

```

for j = 1:nbtokens
    % plotting F0 (raw)
    figure(1)
    F0 = nonzeros(data(:,3,j));
    times = nonzeros(data(:,1,j)) - data(1,1,j);
    plot(times,F0, '-p', 'Color', [(j-1)/nbtokens 0 (1- ((j-1)/nbtokens))], 'MarkerSize',10)
    hold on
    pause(0.9)
    % plotting Oq (raw)
    figure(2)
    nb_periods = length(times);
    for L = 1:nb_periods
        if data(L,10,j) ~= 0
            plot(times(L),data(L,10,j), 'p', 'Color', [(j-1)/nbtokens 0 (1- ((j-1)/nbtokens))], 'MarkerSize',10)
            hold on
        end
    end
end

figure(1)
%setting for axis
axis([-0.01 0.4 0 250])
xlabel('KTM Tone 1: Speaker M1, Experiment 1. Time in seconds.')
ylabel('F_o (Hz)')
orient landscape
%saving as pdf and eps
print('-dpdf', 'KTMM1Exp1Ton1Fo.pdf')
print('-depsec', 'KTMM1Exp1Ton1Fo.eps')

figure(2)
axis([-0.01 0.4 0 90])
xlabel('KTM Tone 1: Speaker M1, Experiment 1. Time in seconds.')
ylabel('O_q (%)')
orient landscape
print('-dpdf', 'KTMM1Exp1Ton1Oq.pdf')
print('-depsec', 'KTMM1Exp1Ton1Oq.eps')

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

%calculation of average curve; 100 points, plus
%syllable simulations, third value means average Fo
%of speaker

```

```
[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
    meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (data1,100,130);
```

```
figure(3)
clf
h = axes;
set(h, 'FontSize',19)

%plotting average F0 with standard derivation
errorbar(ATV,AFV,ASF, '-pb')

axis([-0.01 0.4 0 250])
xlabel('KTM Tone 1: Speaker M1, Experiment 1. Time in
seconds. ')
ylabel('Fo in Hz (Error bars: standard deviation)')
orient landscape
print('-dpdf','KTMM1Exp1Ton1ASF.pdf')
print('-depsec','KTMM1Exp1Ton1ASF.eps')
```

```
figure(4)
clf
h = axes;
set(h, 'FontSize',19)
%plotting average Oq with standard derivation
errorbar(ATV,AQV,ASQ, 'pr')

axis([-0.01 0.4 0 90])
xlabel('KTM Tone 1: Speaker M1, Experiment 1. Time in
seconds. ')
ylabel('O_q in % (Error bars: standard deviation)')
orient landscape
print('-dpdf','KTMM1Exp1Ton1ASQ.pdf')
print('-depsec','KTMM1Exp1Ton1ASQ.eps')
```

9.2.2 For the entire tone system

This script is served for plotting the curves average F_0 or O_q for all tone system in a graph by many ways. Therefore, the output includes 5 graphs:

1. average F_0 curves with standard deviation
2. average O_q curves with standard deviation
3. average F_0 curves in semitone.
4. average F_0 curves with no standard deviation
5. average O_q curves with no standard deviation

Script:

```
% To plot the average F0 and Oq curves of all tone
system in a matrix by many ways
```

```

% clearing workspace
clear

% Indicate the speaker.
spe = 'M1';

% Speaker:
if spe == 'M1'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-
        MTQ_KTM_M1_TONEXP1_REPET_SORTED';
elseif spe == 'M5'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-MTQ_KTM_M5_TONEXP1_SORTED';
elseif spe == 'M6'
    address = 'C:\Users\ASUS\Dropbox\
        MinhChau_research_shared\Mtq_KTM_Study\
        Mtq_MatLab\crdo-MTQ_KTM_M6_TONEXP1_SORTED';
end

% Loading the data
load (address)

figure(1)
% First figure: plotting average F0 with standard
    deviation

% Clearing the figure
clf

%setting a value for font size
h = axes;
set(h, 'FontSize', 19)

%calculation of average F0 curves with standard
    deviation for all tones of tone system
[AFV, AQV, ADV, ADOV, ATV, ASF, ASQ, ASD, ASDO, meanFovector,
    meanOqvector, meanDvector, meanDOvector, timevector] =
    aver (datat1, 100, 130)
errorbar(ATV, AFV, ASF, '-pg');
hold on

[AFV, AQV, ADV, ADOV, ATV, ASF, ASQ, ASD, ASDO, meanFovector,
    meanOqvector, meanDvector, meanDOvector, timevector] =
    aver (datat2, 100, 130)
errorbar(ATV, AFV, ASF, '-sb');

```



```

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
    meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat3,100,130)
errorbar(ATV,AFV,ASF, '-mo');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
    meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat4,100,130)
errorbar(ATV,AFV,ASF, '-r*');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
    meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat5,100,130)
errorbar(ATV,AFV,ASF, '-dk');

%setting for axis
axis([-0.01 0.4 0 250])
hleg1 = legend('tone 1','tone 2','tone 3','tone 4','
    tone 5')
xlabel('The KTM tone system: Speaker M1, Experiment 1.
    Time in seconds.')
ylabel('Fo in Hz (Error bars: standard deviation)')

%saving as pdf and eps
orient landscape
print('-dpdf','KTMM1Exp1_5Tones_ASF.pdf')
print('-depsec','KTMM1Exp1_5Tones_ASF.eps')

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(2)
% Second figure: plotting average Oq with standard
    deviation

clf
h = axes;
set(h,'FontSize',19)

%calculation of average oq curves standard deviation
    for all tones of tone system
[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
    meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat1,100,130)
errorbar(ATV,AQV,ASQ, '-pg');
hold on

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
    meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat2,100,130)
errorbar(ATV,AQV,ASQ, '-sb');

```

```

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
 aver (datat3,100,130)
errorbar(ATV,AQV,ASQ, '-mo');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
 aver (datat4,100,130)
errorbar(ATV,AQV,ASQ, 'r*');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
 aver (datat5,100,130)
errorbar(ATV,AQV,ASQ, '-dk');

axis([-0.01 0.4 0 90])
hleg2 = legend('tone 1','tone 2','tone 3','tone 4','
 tone 5')
xlabel('The KTM tone system: Speaker M1, Experiment 1.
 Time in seconds.')
ylabel('Oq in % (Error bars: standard deviation)')
orient landscape
print('-dpdf','KTMM1Exp1_5Tones_ASQ.pdf')
print('-depsc','KTMM1Exp1_5Tones_ASQ.eps')

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% figure (3)
% Third figure: Plotting average F0 in semitone

clf
h = axes;
set(h,'FontSize',19)

%calculation of average Fo curves in semitone for all
tones of tone system
[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector,
 Oqmatrix,Fomatrix,total,AsemitonesV,ASsemitones,
 Fomatrix_semitones] = aver (datat1,100,112)
errorbar(ATV,AsemitonesV,ASsemitones, '-pg');
hold on

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector,
 Oqmatrix,Fomatrix,total,AsemitonesV,ASsemitones,
 Fomatrix_semitones] = aver (datat2,100,112)
errorbar(ATV,AsemitonesV,ASsemitones, '-sb');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector,

```

```

    Oqmatrix,Fomatrix,total,AsemitonesV,ASsemitones,
    Fomatrix_semitones] = aver (datat3,100,112)
errorbar(ATV,AsemitonesV,ASsemitones, '-mo');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
meanOqvector,meanDvector,meanDOvector,timevector,
Oqmatrix,Fomatrix,total,AsemitonesV,ASsemitones,
Fomatrix_semitones] = aver (datat4,100,112)
errorbar(ATV,AsemitonesV,ASsemitones, '-r*');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
meanOqvector,meanDvector,meanDOvector,timevector,
Oqmatrix,Fomatrix,total,AsemitonesV,ASsemitones,
Fomatrix_semitones] = aver (datat5,100,112)
errorbar(ATV,AsemitonesV,ASsemitones, '-dk');

axis([-0.01 0.4 -20 8])
hleg1 = legend('tone 1','tone 2','tone 3','tone 4','
tone 5')
xlabel('The KTM tone system: Speaker M1, Experiment 1.
Time in seconds.')
ylabel('Fo in semitones (Error bars: standard deviation
)')
orient landscape
print('-dpdf','KTMM1Exp1_5Tones_ASsemitones.pdf')
print('-depsc','KTMM1Exp1_5Tones_ASsemitones.eps')

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(4)
% Fourth figure: Plotting average F0 (non-standard
deviation)

clf
h = axes;
set(h,'FontSize',19)

%calculation of average F0 curves with no standard
deviation for all tones of tone system
[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
meanOqvector,meanDvector,meanDOvector,timevector] =
aver (datat1,100,112)
plot(ATV,AFV, '-pg');
hold on

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
meanOqvector,meanDvector,meanDOvector,timevector] =
aver (datat2,100,112)
plot(ATV,AFV, '-sb');

```

```

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat3,100,112)
plot(ATV,AFV, '-mo');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat4,100,112)
plot(ATV,AFV, '-r*');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat5,100,112)
plot(ATV,AFV, '-dk');

axis([-0.01 0.4 0 250])
hleg3 = legend('tone 1','tone 2','tone 3','tone 4','
    tone 5')
xlabel('The KTM tone system: Speaker M1, Experiment 1.
    Time in seconds.')
ylabel('Fo in Hz (No standard deviation)')
orient landscape
print('-dpdf','KTMM1Exp1_5Tones_AFV.pdf')
print('-depsc','KTMM1Exp1_5Tones_AFV.eps')

% %%%%%%%%%%%
% %%%%%%%%%%%

figure(5)
% Plotting average of Oq (non-standard deviation)

% % Clearing the figure
clf
%
h = axes;
set(h,'FontSize',19)

%calculation of average Oq curves with no standard
    deviation for all tones of tone system
[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat1,100,112)
plot(ATV,AQV, '-pg');
hold on

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
 meanOqvector,meanDvector,meanDOvector,timevector] =
    aver (datat2,100,112)
plot(ATV,AQV, 'sb');

```

```

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
  meanOqvector,meanDvector,meanDOvector,timevector] =
  aver (datat3,100,112)
plot(ATV,AQV, '-mo');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
  meanOqvector,meanDvector,meanDOvector,timevector] =
  aver (datat4,100,112)
plot(ATV,AQV, 'r*');

[AFV,AQV,ADV,ADOV,ATV,ASF,ASQ,ASD,ASDO,meanFovector,
  meanOqvector,meanDvector,meanDOvector,timevector] =
  aver (datat5,100,112)
plot(ATV,AQV, '-dk');

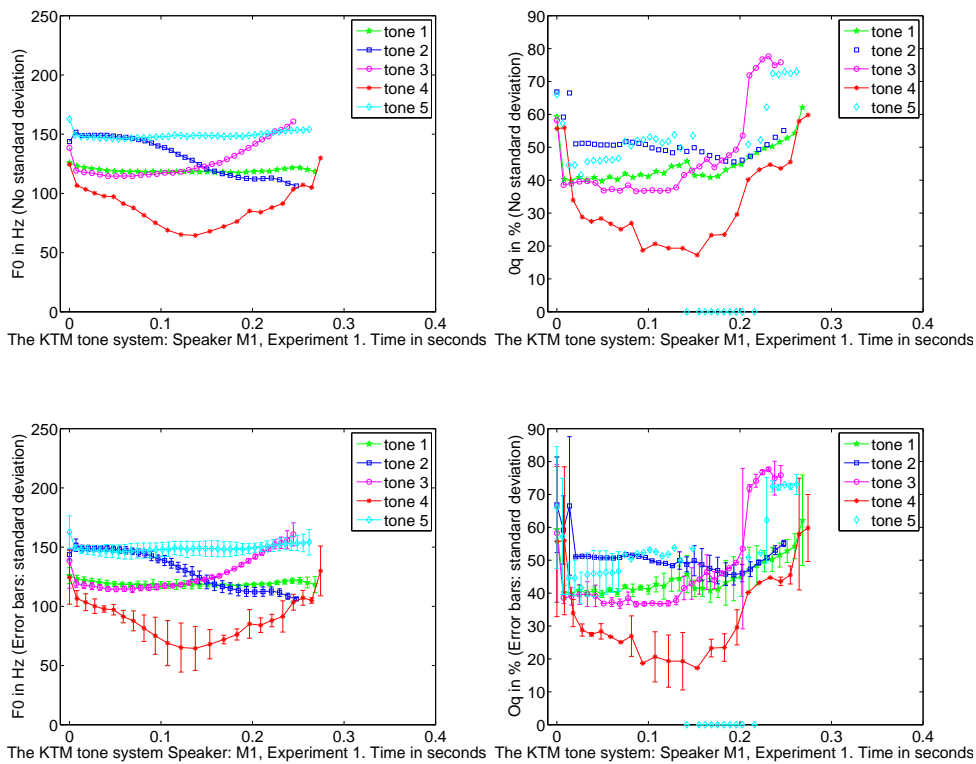
axis([-0.01 0.4 0 90])
hleg4 = legend('tone 1','tone 2','tone 3','tone 4','
  tone 5')
xlabel('The KTM tone system: Speaker M1, Experiment 1.
  Time in seconds.')
ylabel('oq in % (No standard deviation)')
orient landscape
print('-dpdf','KTMM1Exp1_5Tones_AQV.pdf')
print('-depsec','KTMM1Exp1_5Tones_AQV.eps')

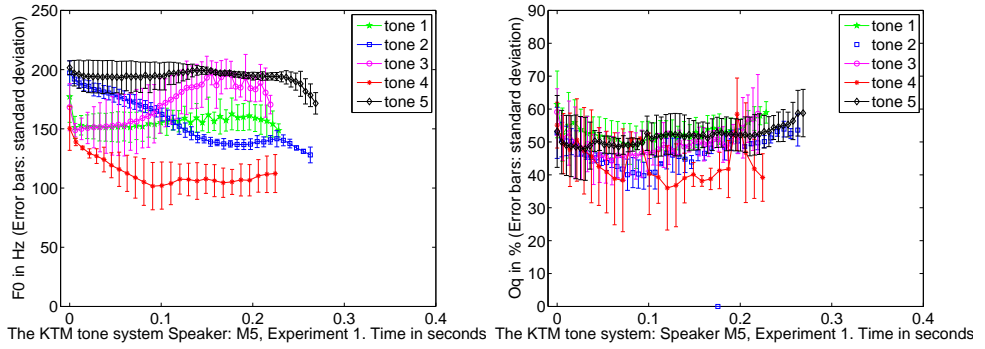
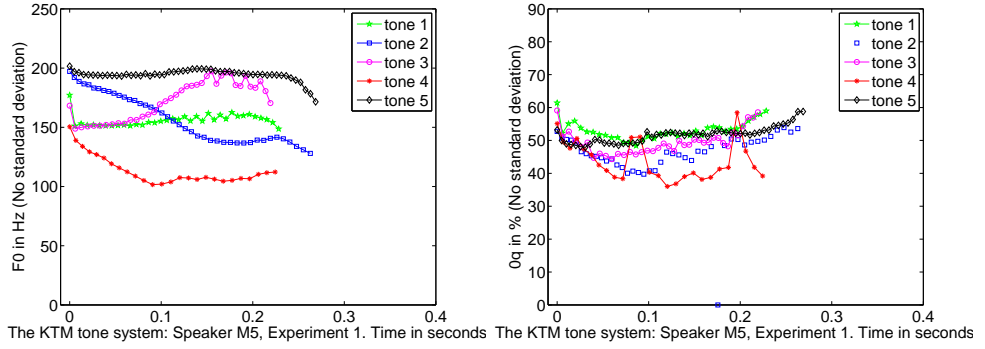
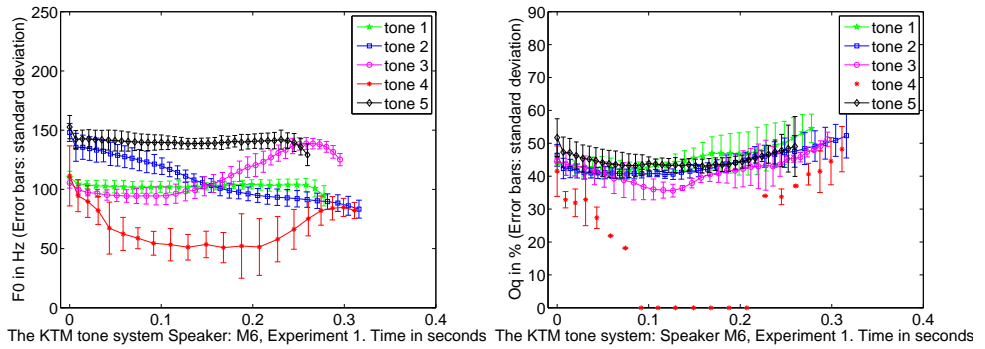
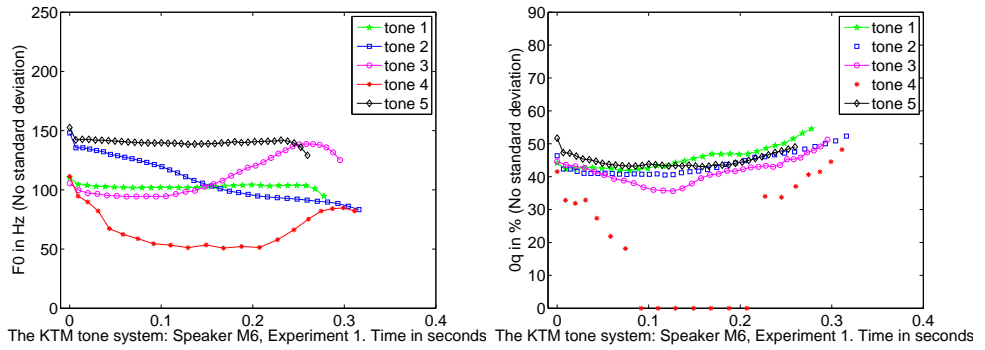
```

APPENDIX 5: ADDITIONAL FIGURES

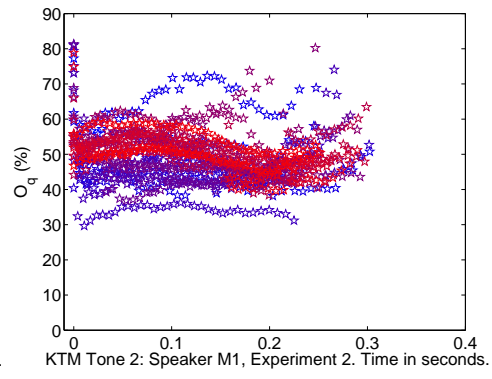
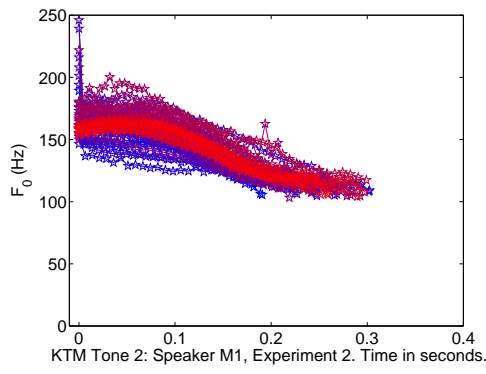
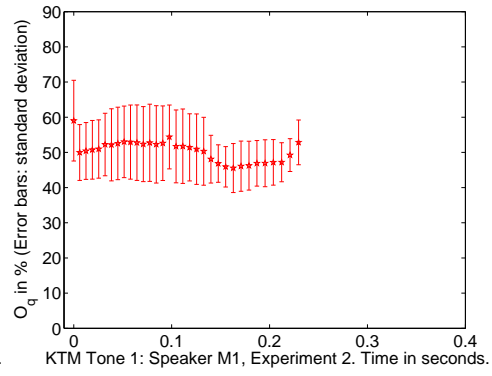
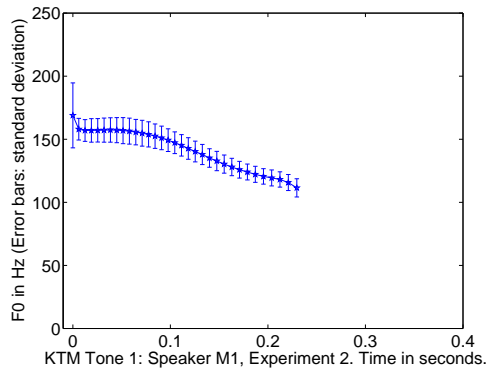
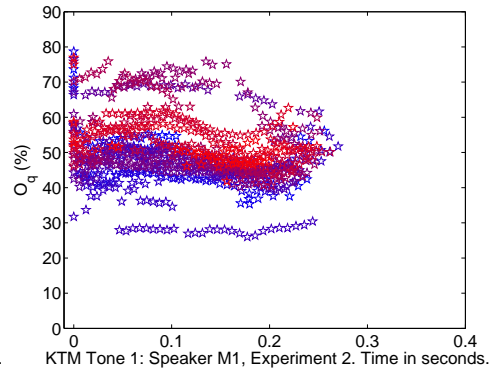
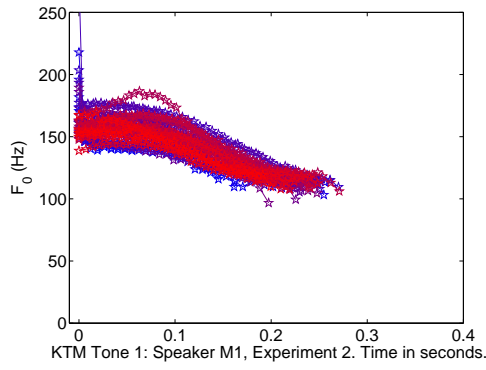
Beside figures used in the analysis and discussion in Chapters 3 and 4, this Appendix provides the rest of experimental analysis results in a systematic way. In other words, a full set of figures will be presented here, as a result of applying MATLAB scripts to recorded materials. It is hoped that this can help readers have a good look on the whole results of this study.

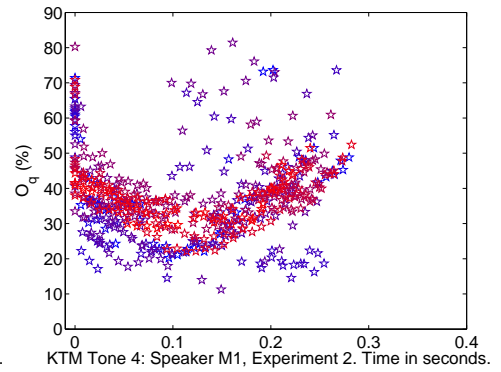
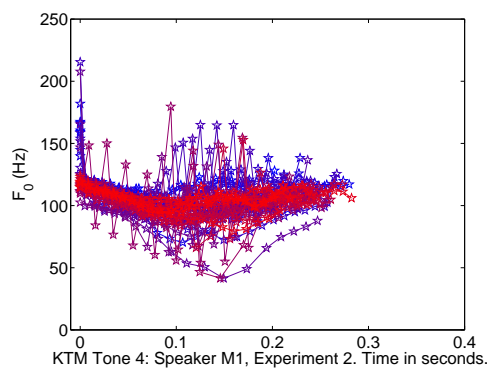
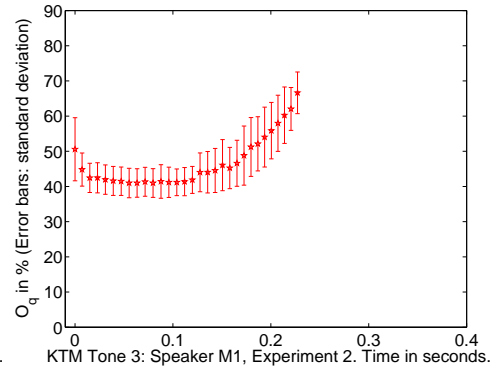
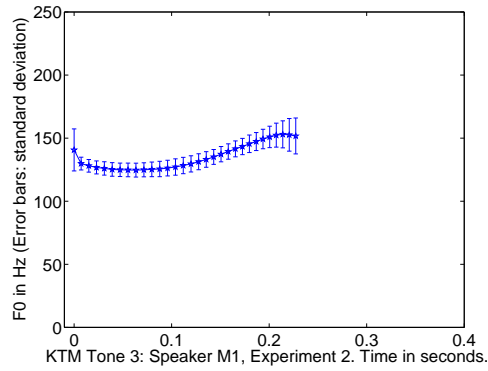
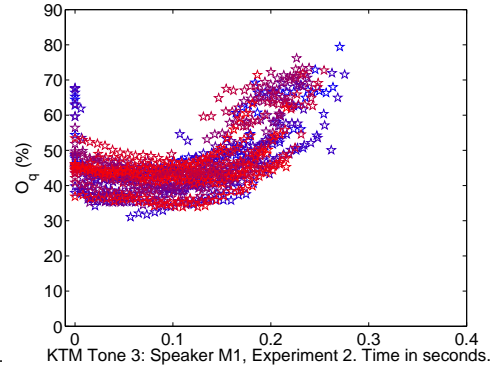
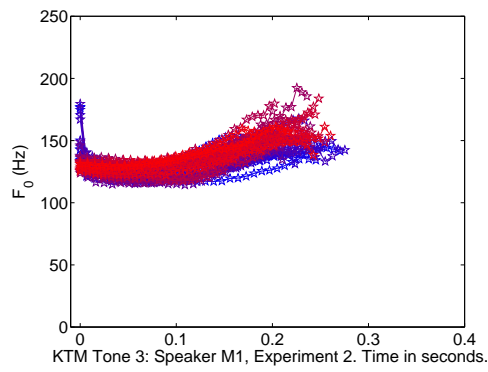
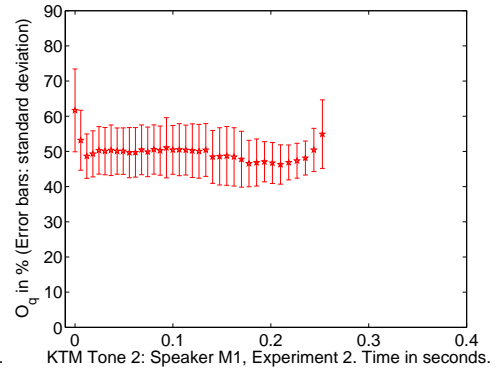
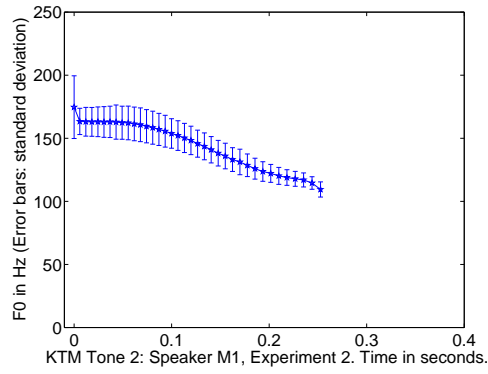
10.1 EXPERIMENT 1

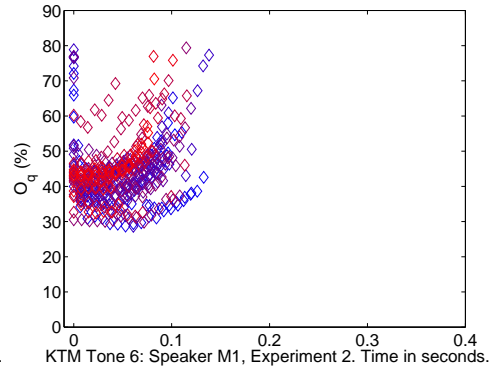
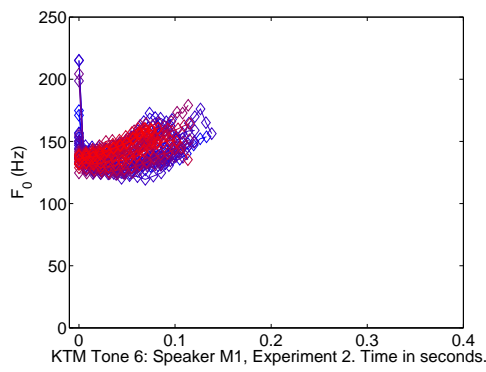
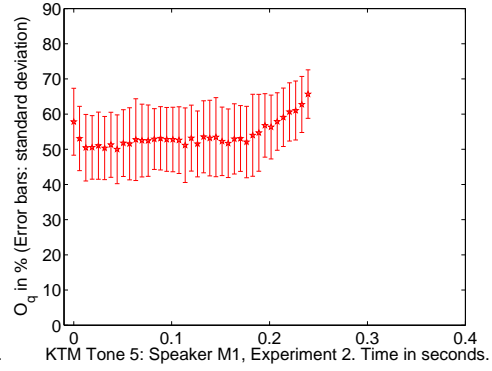
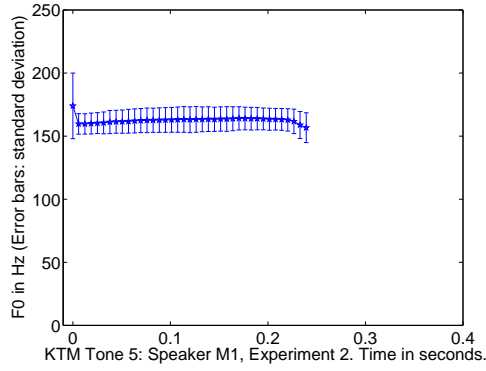
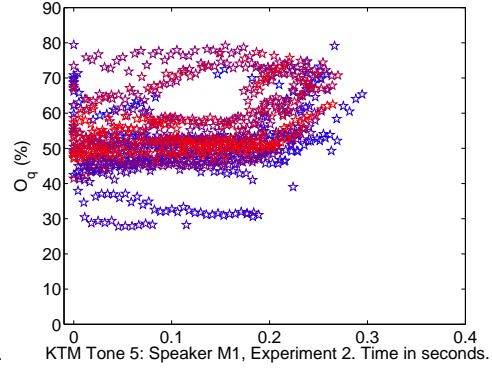
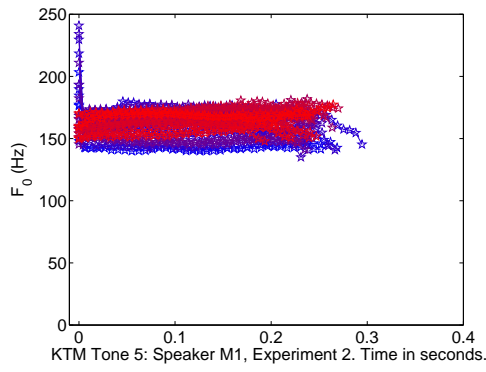
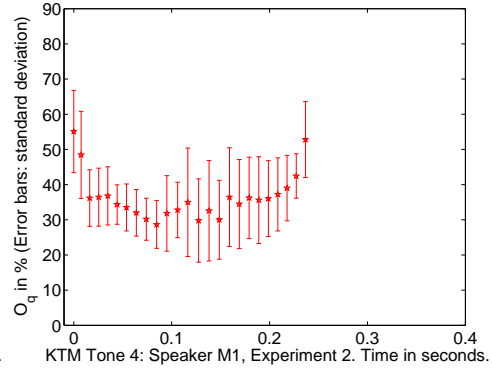
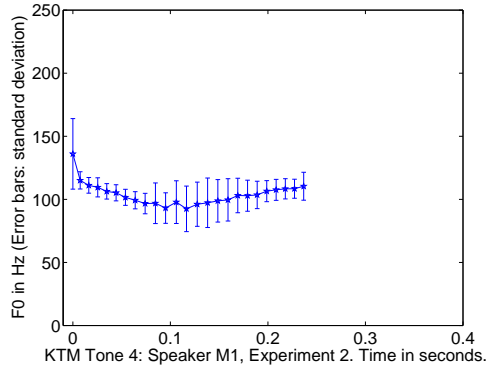
10.1.1 *Speaker M1*

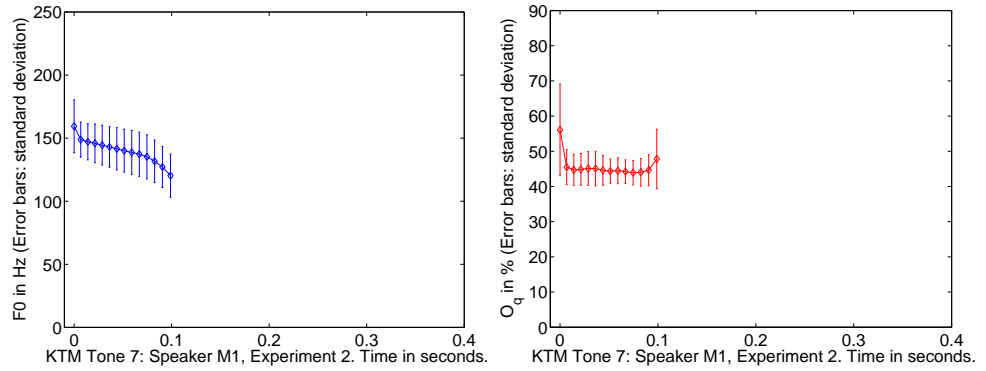
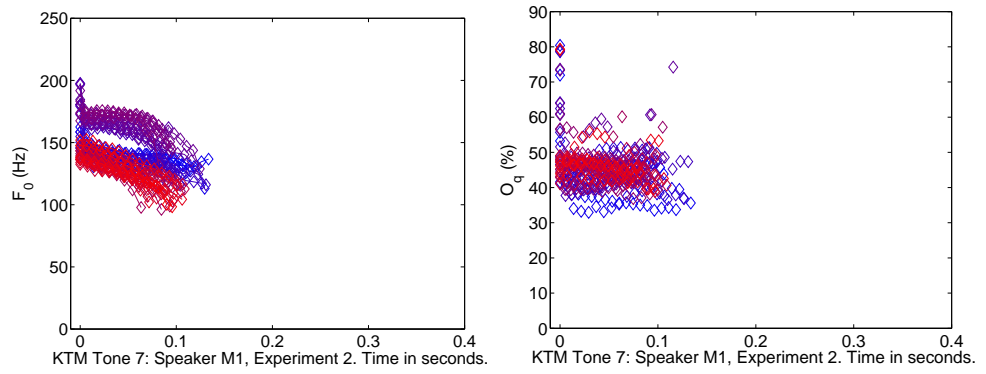
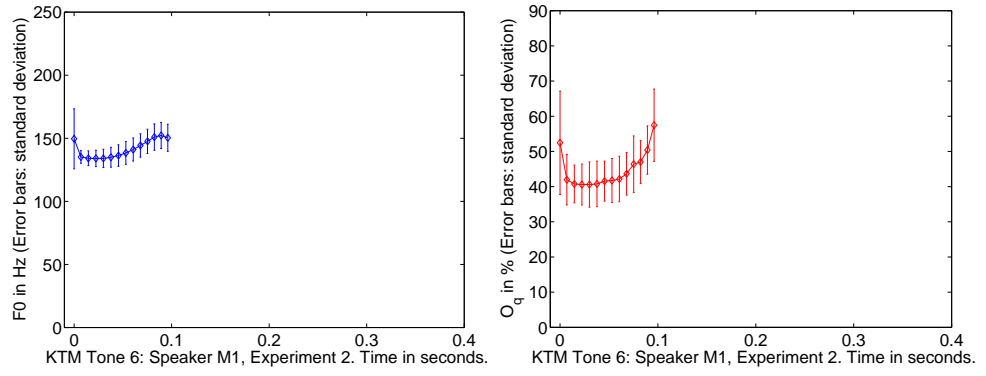
10.1.2 *Speaker M5*10.1.3 *Speaker M6*

10.2 EXPERIMENT 2

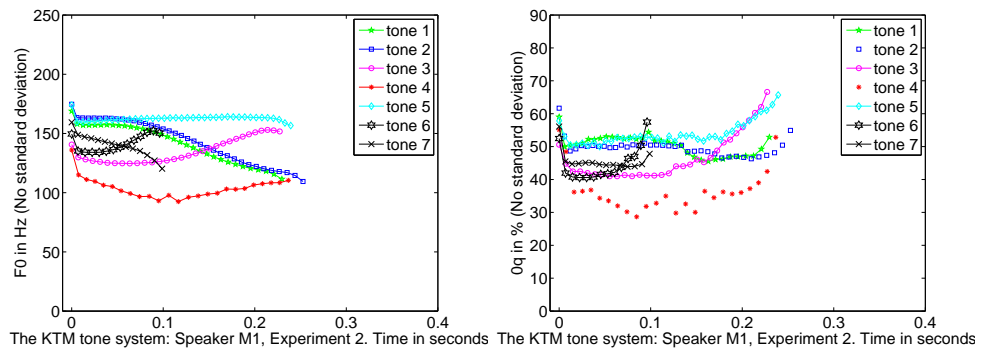
10.2.1 *Each individual tone*

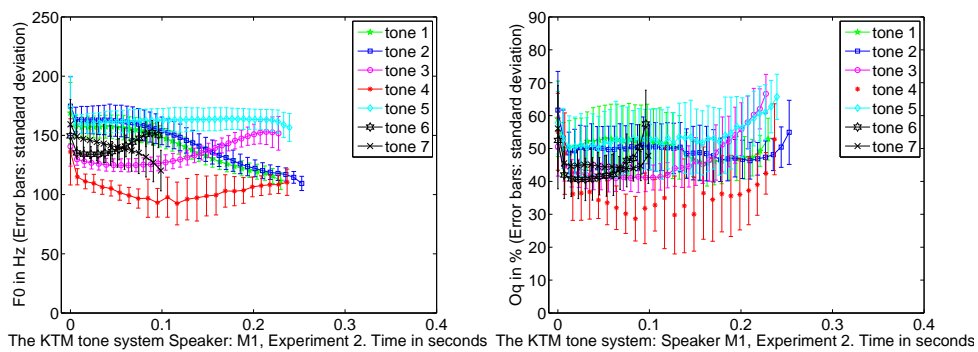






10.2.2 Full tone system





BIBLIOGRAPHY

- [1] Arthur S. Abramson. « The noncategorical perception of tone categories in Thai. » In: *Frontiers of speech communication research*. Ed. by B. Lindblom and S. Öhman. London: Academic Press, 1979, pp. 127–134.
- [2] Arthur S. Abramson, Patrick W Nye, and Therapan Luangthongkum. « Voice register in Khmu: Experiments in production and perception. » In: *Phonetica* 64.2-3 (2007), pp. 80–104.
- [3] Arthur S. Abramson, Therapan L. Thongkum, and Patrick W. Nye. « Voice register in Suai (Kuai): An analysis of perceptual and acoustic data. » In: *Phonetica* 61.2-3 (2004), pp. 147–171.
- [4] Arthur S. Abramson and Kalaya Tingsabadh. « Thai Final Stops: Cross-Language Perception. » In: *Phonetica* 56.3-4 (1999), pp. 111–122.
- [5] Milton E. Barker and Muriel A. Barker. *Mường-Vietnamese-English dictionary*. Huntington Beach, CA: Summer Institute of Linguistics, 1976, 537 pp.
- [6] Miriam A. Barker. « Bibliography of Mường and other Vietic language groups, with notes. » In: *The Mon-Khmer Studies Journal* 23 (1993), pp. 197–243. URL: <http://purl.org/sealang/barker1993bibliography.pdf>.
- [7] Alexander Melville Bell. *Visible Speech: The Science of Universal Alphabets Or, Self-interpreting Physiological Letters, for the Writing of All Languages in One Alphabet*. Simpkin, Marshall & Company, 1867.
- [8] Thiện Bùi. *Dân ca Mường: Phần tiếng Mường*. NXB Văn hóa Dân tộc, 2010.
- [9] Marc Brunelle. « Northern and Southern Vietnamese tone coarticulation: A comparative case study. » In: *Journal of Southeast Asian Linguistics* 1 (2009), pp. 49–62.
- [10] Marc Brunelle. « Tone perception in Northern and Southern Vietnamese. » In: *Journal of Phonetics* 37 (2009), pp. 79–96.
- [11] Marc Brunelle. « Dialect experience and perceptual integrality in phonological registers: Fundamental frequency, voice quality and the first formant in Cham. » In: *Journal of the Acoustical Society of America* 131.4 (2012), pp. 3088–3102.
- [12] Marc Brunelle. « Effects of lexical frequency and lexical category on the duration of Vietnamese syllables. » In: *Proceedings of ICPhS XVIII*. Glasgow, 2015.
- [13] Marc Brunelle and Joshua Finkeldey. « Tone perception in Sgaw Karen. » In: *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 17)*. 2011, pp. 372–375.

- [14] Marc Brunelle, Kiều Phương Hạ, and Martine Grice. « Intonation in Northern Vietnamese. » In: *The Linguistic Review* 29.1 (2012), pp. 3–36.
- [15] Marc Brunelle, Nguyễn Khắc Hùng, and Nguyễn Duy Dương. « A laryngographic and laryngoscopic study of Northern Vietnamese tones. » In: *Phonetica* 67.3 (2010), pp. 147–169.
- [16] Marc Brunelle and Stefanie Jannedy. « Social effects on the perception of Vietnamese tones. » In: *International Congress of Phonetic Sciences*. Saarbrücken, 2007, pp. 1461–1464.
- [17] Marc Brunelle and James Kirby. « Re-assessing tonal diversity and geographical convergence in Mainland Southeast Asia. » In: *Languages of Mainland Southeast Asia: The State of the Art* 649 (2015), pp. 82–111.
- [18] Marc Brunelle and James Kirby. « Tone and phonation in Southeast Asian languages. » In: *Language and Linguistics Compass* 10.4 (2016), pp. 191–207.
- [19] Denis Burnham and Elizabeth Francis. « The role of linguistic experience in the perception of Thai tones. » In: *Southeast Asian linguistic studies in honour of Vichin Panupong* (1997), pp. 29–47.
- [20] D.G. Childers and C.K. Lee. « Vocal quality factors: Analysis, synthesis and perception. » In: *Journal of the Acoustical Society of America* 90.5 (1991), pp. 2394–2410.
- [21] Donald G Childers and Ashok K Krishnamurthy. « A critical review of electroglottography. » In: *Critical reviews in biomedical engineering* 12.2 (1984), pp. 131–161.
- [22] Paul Christophersen. « The glottal stop in English. » In: (1952). URL: <http://www.tandfonline.com/doi/pdf/10.1080/00138385208596879> (visited on 10/04/2016).
- [23] Nick Clements, Alexis Michaud, and Cédric Patin. « Do we need tone features? » In: *Tones and Features*. Ed. by Elizabeth Hume, John Goldsmith, and W. Leo Wetzels. Berlin: De Gruyter Mouton, 2011, pp. 3–24.
- [24] Raymond H. Colton and Edward G. Conture. « Problems and Pitfalls of Electroglottography. » In: *Journal of Voice* 4.1 (1990), pp. 10–24.
- [25] David Crystal. *Dictionary of linguistics and phonetics*. Vol. 30. John Wiley & Sons, 2011.
- [26] Jeanne Cuisinier. *Les Mu'ò'ng: Géographie humaine et sociologie*. Vol. 45. Institut d'ethnologie, 1948.
- [27] Gilles Degottex, John Kane, Thomas Drugman, Tuomo Raitio, and Stefan Scherer. « COVAREP: a collaborative voice analysis repository for speech technologies. » In: *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. DOI 10.1109/ICASSP.2014.6853739. IEEE, 2014, pp. 960–964. URL: <https://github.com/covarep/covarep>.

- [28] Trần Trí Dõi. *Nghiên cứu ngôn ngữ các dân tộc thiểu số ở Việt Nam* [*Linguistic research about the ethnic minorities of Vietnam*]. NXB Đại học Quốc gia Hà Nội, 2000.
- [29] John H Esling. « The IPA Categories “Pharyngeal” and “Epiglottal” Laryngoscopic Observations of Pharyngeal Articulations and Larynx Height. » In: *Language and Speech* 42.4 (1999), pp. 349–372.
- [30] John H Esling, Katherine E Fraser, and Jimmy G Harris. « Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). » In: *Journal of Phonetics* 33.4 (2005), pp. 383–410.
- [31] John Esling. « Laryngographic study of phonation type and laryngeal configuration. » In: *Journal of the International Phonetic Association* 14 (1984), pp. 56–73.
- [32] John Esling. « Pharyngeal consonants and the aryepiglottic sphincter. » In: *Journal of the International Phonetic Association* 26.02 (1996), pp. 65–88.
- [33] Philippe Fabre. « Un procédé électrique percutané d’inscription de l’accolement glottique au cours de la phonation: glottographie de haute fréquence. » In: *Bulletin de l’Académie Nationale de Médecine* 141 (1957), pp. 66–69.
- [34] Philippe Fabre. « Etude comparée des glottogrammes et des phonogrammes de la voix humaine. » In: *Annuaire Oto-rhino Laryngologie* 75 (1958), pp. 767–775.
- [35] Philippe Fabre. « La glottographie électrique en haute fréquence: Particularités de l’appareillage. » In: *Comptes rendus des séances de la Société de biologie: et de ses filiales* 153.8-9 (1959), pp. 1361–1364.
- [36] Philippe Fabre. « Glottographie respiratoire, appareillage et premiers résultats. » In: *Comptes rendus hebdomadaires des séances* 252.9 (1961), p. 1386.
- [37] Michel Ferlus. « Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien. » In: *Cahiers de linguistique - Asie Orientale* 11.1 (1982), pp. 83–106.
- [38] Michel Ferlus. « Langues et peuples viet-muong. » In: *Mon-Khmer Studies* 26 (1996), pp. 7–28.
- [39] Michel Ferlus. « Les systèmes de tons dans les langues viet-muong. » In: *Diachronica* 15.1 (1998), pp. 1–27.
- [40] Michel Ferlus. « Les disharmonies tonales en viet-muong et leurs implications historiques [Irregular tonal correspondences within Vietic and their historical implications]. » In: *Cahiers de Linguistique - Asie Orientale* 28.1 (1999), pp. 83–99.

- [41] Michel Ferlus. « Les hypercorrections dans le thổ de Làng Lỗ (Nghệ An, Vietnam) ou les pièges du comparatisme. » In: *Quinzièmes Journées de Linguistique de l'Asie Orientale*. Ecole des Hautes Etudes en Sciences Sociales, Paris, 2001. URL: <http://halshs.archives-ouvertes.fr/halshs-00922722>.
- [42] Christer Gobl. « Voice Source Variation and Its Communicative Functions. » In: *The handbook of phonetic sciences* (2010), pp. 378–423.
- [43] Matthew Gordon and Peter Ladefoged. « Phonation types: a cross-linguistic overview. » In: *Journal of Phonetics* 29 (2001), pp. 383–406.
- [44] James Frederick Gruber. « An articulatory, acoustic, and auditory study of Burmese tone. » In: (2011). URL: <https://repository.library.georgetown.edu/handle/10822/558130> (visited on 05/23/2016).
- [45] Mieko S. Han and Kong-On Kim. « Phonetic variation of Vietnamese tones in disyllabic utterances. » In: *Journal of Phonetics* 2 (1974), pp. 223–232.
- [46] Jimmy G Harris. « States of the glottis for voiceless plosives. » In: *Proceedings of the 14th international congress of phonetic sciences*. Vol. 3. 1999, pp. 2041–2044.
- [47] Jimmy G. Harris. « States of the glottis of Thai voiceless stops and affricates. » In: *Essays in Tai linguistics*. Ed. by Kalaya Tingsabadh and Arthur S. Abramson. Bangkok: Chulalongkorn University Press, 2001, pp. 3–11.
- [48] André-Georges Haudricourt. « Les voyelles brèves du vietnamien. » In: *Bulletin de la Société de Linguistique de Paris* 48.1 (1952), pp. 90–93.
- [49] André-Georges Haudricourt. « La place du vietnamien dans les langues austroasiatiques. » In: *Bulletin de la Société de Linguistique de Paris* 49.1 (1953), pp. 122–128.
- [50] André-Georges Haudricourt. « Comment reconstruire le chinois archaïque. » In: *Word* 10.2-3 (1954), pp. 351–364.
- [51] André-Georges Haudricourt. « De l'origine des tons en vietnamien. » In: *Journal Asiatique* 242 (1954), pp. 69–82.
- [52] La Vaughn H. Hayes. « Vietic and Việt-Mường: a new subgrouping in Mon-Khmer. » In: *Mon-Khmer Studies* 21 (1992), pp. 211–228.
- [53] Roe-Merrill S. Heffner. *General Phonetics*. Madison: The University of Wisconsin Press, 1950.
- [54] Nathalie Henrich, Cédric Gendrot, and Alexis Michaud. *Tools for Electrolottographic Analysis: Software, Documentation and Databases*. 2005. URL: <http://voiceresearch.free.fr/egg/index.html>.

- [55] Nathalie Henrich, Christophe d'Alessandro, Michèle Castellengo, and Boris Doval. « On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. » In: *Journal of the Acoustical Society of America* 115.3 (2004), pp. 1321–1332.
- [56] William Holder. *Elements of Speech: An Essay of Inquiry into the Natural Production of Letters: With an Appendix Concerning Persons Deaf & Dumb*. 1669.
- [57] Jean-Marie Hombert. « Consonant types, vowel quality and tone. » In: *Tone : a Linguistic Survey*. Ed. by Victoria A. Fromkin. New York: Academic Press, 1978, pp. 77–111.
- [58] Daniel Jones. *Outline of English phonetics (8th ed.)* New York: E.P. Dutton & Co, 1918.
- [59] Patricia Keating, Marc Garellek, and Jody Kreiman. « Acoustic properties of different kinds of creaky voice. » In: *ICPhS XVIII (18th International Congress of Phonetic Sciences)*. 2015.
- [60] James Kirby. « Dialect experience in Vietnamese tone perception. » In: *Journal of the Acoustical Society of America* 127.6 (2010), pp. 3749–3757.
- [61] James Kirby. « Vietnamese (Hanoi Vietnamese). » In: *Journal of the International Phonetic Association* 41.3 (2011), pp. 381–392.
- [62] James Kirby. « Incipient tonogenesis in Phnom Penh Khmer: Acoustic and perceptual studies. » In: *Journal of Phonetics* 43 (2014), pp. 69–85.
- [63] Paul L Kirk, Jenny Ladefoged, and Peter Ladefoged. « Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. » In: *American Indian linguistics and ethnography in honor of Laurence C. Thompson* (1993), pp. 435–450.
- [64] Dennis Klatt and Laura Klatt. « Analysis, synthesis, and perception of voice quality variations among female and male talkers. » In: *Journal of the Acoustical Society of America* 87 (1990), pp. 820–857.
- [65] Klaus J. Kohler. « Glottal stops and glottalization in German. » In: *Phonetica* 51 (1994), pp. 38–51.
- [66] Jianjing Kuang. « Phonation in Tonal Contrasts. » Ph. D. thesis. PhD thesis. Los Angeles: University of California, 2013.
- [67] Jianjing Kuang. « The tonal space of contrastive five level tones. » In: *Phonetica* 70.1-2 (2013), pp. 1–23. DOI: [10.1159/000353853](https://doi.org/10.1159/000353853).
- [68] William Labov. *Principles of linguistic change. Internal factors*. Language in Society 20. Oxford: Basil Blackwell, 1994.
- [69] William Labov. *Principles of linguistic change. Social factors*. Language in Society 29. Oxford: Basil Blackwell, 2001.
- [70] William Labov. *Principles of linguistic change. Volume 3: Cognitive and cultural factors*. Language in Society 39. Oxford, UK & Malden, USA: Wiley-Blackwell, 2010.

- [71] Peter Ladefoged. *Preliminaries to linguistic phonetics*. Chicago & London: University of Chicago Press, 1971.
- [72] Peter Ladefoged. *A Course in Phonetics*. Harcourt Brace College Publishers, 1975.
- [73] Peter Ladefoged. « The linguistic use of different phonation types. » In: *Vocal fold physiology: Contemporary research and clinical issues* (1983), pp. 351–360.
- [74] Peter Ladefoged and Ian Maddieson. *The sounds of the world's languages*. Ed. by M. Kenstowicz, J. Goldsmith, Nick Clements, and D. Steriade. Phonological Theory. Oxford, U.K. & Cambridge, Massachusetts: Blackwell, 1996.
- [75] Peter Ladefoged, Ian Maddieson, and M. Jackson. « Investigating phonation types in different languages. » In: *Vocal fold physiology: voice production, mechanisms and functions*. Ed. by Osamu Fujimura. New York: Raven Press, 1988, pp. 297–317.
- [76] John Laver. *The phonetic description of voice quality*. Cambridge: Cambridge University Press, 1980.
- [77] John Laver. *Principles of Phonetics*. Cambridge: Cambridge University Press, 1994.
- [78] M Paul Lewis, Gary F Simons, and Charles D Fennig. *Ethnologue: Languages of the world*. Vol. 9. SIL international Dallas, TX, 2009.
- [79] Timothy Light. « Tonogenesis: analysis and implications. » In: *Lingua* 46.2-3 (1978), pp. 115–131.
- [80] B. Lindblom. « Explaining phonetic variation: a sketch of the H&H theory. » In: *Speech production and speech modelling*. Ed. by W.J. Hardcastle and Alain Marchal. Dordrecht: Kluwer, 1990, pp. 403–439.
- [81] Nguyễn Van Loi and Jerold A. Edmondson. « Tones and voice quality in modern northern Vietnamese: instrumental case studies. » In: *Mon-Khmer Studies* 28 (1998), pp. 1–18.
- [82] Dang-Khoa Mac, Thi-Lan Nguyen, Alexis Michaud, and Do-Dat Tran. « Influences of speaker attitudes on glottalized tones: a study of two Vietnamese sentence-final particles. » In: *Proceedings of ICPhS XVIII (18th International Congress of Phonetic Sciences)*. Areas: (i) Tone; (ii) SpeechProsody; (iii) Phonation and Voice Quality. Glasgow, 2015.
- [83] André Malécot. « The glottal stop in French. » In: *Phonetica* 31.1 (1975), pp. 51–63.
- [84] André Martinet. « Fricatives and spirants. » In: *Suniti Kumar Chatterji commemoration volume*. Ed. by Bhakti Prasad Mallik. Burdwan, West Bengal, India: Burdwan University Press, 1981, pp. 145–151.
- [85] Henri Maspero. « Etude sur la phonétique historique de la langue annamite: les initiales. » In: *Bulletin de l'Ecole Française d'Extrême-Orient* 12 (1912), pp. 1–127.

- [86] Martine Mazaudon and Alexis Michaud. « Tonal contrasts and initial consonants: a case study of Tamang, a ‘missing link’ in tonogenesis. » In: *Phonetica* 65.4 (2008), pp. 231–256.
- [87] Boyd Michailovsky. *La langue hayu*. Ed. by Sylvain Auroux. Sciences du langage. Paris: CNRS Editions, 1988.
- [88] Boyd Michailovsky, Martine Mazaudon, Alexis Michaud, Séverine Guillaume, Alexandre François, and Evangelia Adamou. « Documenting and researching endangered languages: the Pangloss Collection. » In: *Language Documentation and Conservation* 8 (2014), pp. 119–135. ISSN: 1934-5275. URL: <http://hdl.handle.net/10125/4621>.
- [89] Alexis Michaud. « A Measurement from Electroglossography: DECPA, and its Application in Prosody. » In: *Speech Prosody 2004*. Ed. by Bernard Bel and Isabelle Marlien. Nara, Japan, 2004, pp. 633–636.
- [90] Alexis Michaud. « Final consonants and glottalization: new perspectives from Hanoi Vietnamese. » In: *Phonetica* 61.2-3 (2004), pp. 119–146.
- [91] Alexis Michaud. « Monosyllabicization: patterns of evolution in Asian languages. » In: *Monosyllables: from phonology to typology*. Ed. by Nicole Nau, Thomas Stolz, and Cornelia Stroh. Berlin: Akademie Verlag, 2012, pp. 115–130. URL: <http://halshs.archives-ouvertes.fr/halshs-00436432/>.
- [92] Alexis Michaud, Jacqueline Vaissière, and Minh-Châu Nguyễn. « Phonetic insights into a simple level-tone system: ‘careful’ vs. ‘impatient’ realizations of Naxi High, Mid and Low tones. » In: *ICPhS XVIII (18th International Congress of Phonetic Sciences)*. 2015.
- [93] Alexis Michaud and Tuan Vu-Ngoc. « Glottalized and nonglottalized tones under emphasis: open quotient curves remain stable, Fo curve is modified. » In: *Speech Prosody 2004*. Ed. by Bernard Bel and Isabelle Marlien. Nara, Japan, 2004, pp. 745–748.
- [94] Thi-Lan Nguyen, Alexis Michaud, Do-Dat Tran, and Dang-Khoa Mac. « The interplay of intonation and complex lexical tones: how speaker attitudes affect the realization of glottalization on Vietnamese sentence-final particles. » In: *Proceedings of Interspeech 2013*. Lyon, 2013.
- [95] Oliver Niebuhr. « Perceptual Study of Timing Variables in Fo Peaks. » In: *15th International Congress of Phonetic Sciences*. Barcelona, 2003, pp. 1225–1218.
- [96] Oliver Niebuhr and Alexis Michaud. « Speech data acquisition: the underestimated challenge. » In: *KALIPHO - Kieler Arbeiten zur Linguistik und Phonetik* 3 (2015), pp. 1–42.
- [97] George Noël-Armfield. *General phonetics: For missionaries and students of languages (4th ed)*. W. Heffer & sons Limited, 1931.

- [98] Robert F. Orlikoff. « Scrambled EGG: The Uses and Abuses of Electroglossography. » In: *Phonoscope* 1.1 (1998), pp. 37–53.
- [99] Andrea Hoa Pham. « Vietnamese tone: Tone is not pitch. » Ph. D. dissertation. PhD thesis. Université de Toronto, 2001.
- [100] John Phan. « Mường is not a subgroup: phonological evidence for a paraphyletic taxon in the Viet-Muong sub-family. » In: *Mon-Khmer Studies* 40 (2012), pp. 1–18.
- [101] Kenneth L. Pike. *Tone Languages. A Technique for Determining the Number and Type of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion*. Ann Arbor: University of Michigan Press, 1948.
- [102] Mark Post. « Compounding and the structure of the Tani lexicon. » In: *Linguistics of the Tibeto-Burman Area* 29.1 (2006), pp. 41–60.
- [103] Joanna Przedlacka. « Estuary English: glottaling in the Home Counties. » In: *Oxford University Working Papers in Linguistics, Philology and Phonetics* 5 (2000), pp. 19–24.
- [104] Alexandre de Rhodes. *Dictionarium Annamiticum Lusitanum et Latinum*. Rome, 1651.
- [105] Bernard Roubeau, Nathalie Henrich, and Michèle Castellengo. « Laryngeal vibratory mechanisms: the notion of vocal register revisited. » In: *Journal of Voice* 23.4 (2009), pp. 425–38.
- [106] Richard Keith Sprigg. « The glottal stop and glottal constriction in Lepcha, and borrowing from Tibetan. » In: *Bulletin of Tibetology* 3.1 (1966), pp. 5–14.
- [107] Johan Sundberg. « Maximum speed of pitch changes in singers and untrained subjects. » In: *Journal of Phonetics* 7 (1979), pp. 71–79.
- [108] Jan-Olof Svantesson and David House. « Tone production, tone perception and Kammu tonogenesis. » In: *Phonology* 23 (2006), pp. 309–333.
- [109] Henry Sweet. *A handbook of phonetics: including a popular exposition of the principles of spelling reform*. Cambridge University Press, 2013.
- [110] Trần Tù. « Người Mường ở Hòa Bình cũ. » In: *Người Mường với văn hóa cổ truyền Mường Bì*. Sở Văn hóa Thông tin Hà Sơn Bình, 1988.
- [111] Nguyễn Văn Tài. *Ngữ âm tiếng Mường qua các phương ngữ [The phonetics of the Mường language across its various dialects]*. Hanoi: Nhà Xuất bản Từ điển Bách khoa, 2005.
- [112] Elmar Ternes. *The phonemic analysis of Scottish Gaelic, based on the dialect of Applecross, Ross-shire*. Vol. Third revised edition, with an additional chapter. Dublin: School of Celtic Studies of the Dublin Institute for Advanced Studies, 2006.
- [113] Uriel Weinreich. « On the description of phonic interference. » In: *Word* 13.1 (1957), pp. 1–11.

- [114] Uriel Weinreich. *Languages in contact: French, German and Romansh in twentieth-century Switzerland*. John Benjamins Publishing, 2011.
- [115] Uriel Weinreich, William Labov, and Marvin R Herzog. *Empirical foundations for a theory of language change*. Austin: University of Texas Press, 1968.
- [116] Dough H. Whalen and Andrea G. Levitt. « The universality of intrinsic Fo of vowels. » In: *Journal of Phonetics* 23 (1995), pp. 349–366.
- [117] Bodo Winter. « The other N: The role of repetitions and items in the design of phonetic experiments. » In: *Proceedings of the 18th International Congress of Phonetics Sciences (ICPhS2015)*. Secc Glasgow Scotland UK, 2015.
- [118] Elizabeth Zsiga and Rattima Nitisaroj. « Tone features, tone perception, and peak alignment in Thai. » In: *Language and Speech* 50.3 (2007), pp. 343–383.

DECLARATION

I hereby certify that all work presented in this dissertation is my own unless otherwise acknowledged. The presented results are truthful and have not been published in any other person's work.

Hanoi, Vietnam, 2016

Minh-Châu Nguyễn