

UNIVERSITI TEKNOLOGI MARA

**PROSODIC CONTOURS FOR
CLASSIFICATION OF PERCEIVED
BOUNDARIES ON MALAY
PHRASING**

**HASLIZATUL FAIRUZ BINTI MOHAMED
HANUM**

Thesis submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Information Technology and Quantitative Science)

Faculty of Computer and Mathematical Sciences

October 2021

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the result of my own work unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student : Haslizatul Fairuz Mohamed Hanum

Student I.D. No. : 2011218432

Programme : Doctor of Philosophy (Information Technology and Quantitative Science) – CS990

Faculty : Computer and Mathematical Sciences

Thesis : Prosodic Contours for Classification of Perceived Boundaries on Malay Phrasing



Signature of Student :

Date : October 2021

ABSTRACT

Speech is structured in cohesive sequences of information units defined with unique speaker-related intonation patterns, known as prosody. To capture the communicative intention of a speaker, listeners first focus on the change of rising and/or falling prosody patterns (known as speech events) on words and phrasing to capture the degree of information conveyed. Secondly, intonational breaks are used by speakers to convey speech content in a sequence of smaller sub-segments known as phrasing. Incorrect perception of the events and boundaries causes the communicative intention to be falsely understood by listeners. Thus, previous works evaluate the role of boundaries in languages like English, Swedish, Japanese, and Chinese to help automatic speech recognition and understanding domain. The main issue addressed by previous work on boundary classification is the disagreement on prosody event labelling. Boundaries are falsely classified due to limitations on expert evaluation of prosody events and lack of a prosody standard. Also, evaluating how the pause and intonation variations discriminate between a phrase with a continual content from other subsequent phrases that carry the end (final) content is still an open question. Limited work on the role of prosody is often defined based on specific language lexical rules, which degrades understanding of phrasing and speech content for the under-resourced language. As related research are limited to prosody stress indicating prominence on individual Malay words; thus, processes for defining the prosody governing Malay phrase boundaries are essential. This research aims to evaluate the correlates of prosody on boundaries with a deep (finality) or a shallow (continual) content. Dataset for the boundary classification task contains formal statements recorded from the debate session from the Malaysian Parliamentary Speech recordings. In the first phase, a refined set of Rapid Phrase Prosody Tasks (RPIT) instructions labelled the speech signal with perceived continual or final content. Responses from four male and four female volunteers are analysed using KAPPA and Krippendorff analysis to construct a Malay phrase (MySP) boundaries dataset with an average of 85% agreement on perceived boundary labels from the first research phase. Pitch regression and the rise-fall-connection (RFC) contour parameters are used to model the role of prosody on the perceived boundaries. A new phrase strength correlates computed as slope feature from the highest nuclei point towards the end of the boundary word region. The role of each vector combined with the word and silence durations to signify each deep (final) and shallow (continual) boundary is tested with supervised classifiers. The supervised K-Nearest Neighbour (KNN), Random Forest (RF), and Logistic Regression (Log-Reg) models predicted the boundary classes with up to 75% accuracy. A higher degree of slope excursion is observed on boundaries that the listeners perceived as the deep boundaries (evaluated with finality in speech content through the RPIT) and used to improve classification results on up to 20% of the falsely classified boundaries. This study contributes to the Malay prosody knowledge and classification of phrasing the boundaries.

ACKNOWLEDGEMENT

I am always grateful for the opportunity to embark on this Ph.D. work which allowed me to grow and shrink dynamically with each challenging experience. Special thanks to the supervisor, Associate Professor Ts. Dr. Nur Atiqah Sia Abdullah, for sharing the challenges and incessant work-study provision through this journey. Also, most profound appreciation to the supervisors, Professor Dr. Zainab Abu Bakar FASc, and Dr. Nuru'l 'Izzah Othman, for the persistent trust to let me reach the stars while keeping me grounded.

So many humble people have passionately involved and contributed to this journey, and I hope many best things will continue to happen to each of you, as they are blessed on me. Special love to the marvellous progenies at home, heartening families, and beautiful friends for the companionship and du'a through tussles.

In every fight and procrastination, each grin and smirk assured that the best is yet to come. Much love for the members of '*kite2 aje*'. May each of us is blessed with happiness for lovingly watched over me through the blues and greys.

Thank you to Malaysian Public Services (JPA) for the sponsorship with Universiti Teknologi MARA (UiTM) for encouraging the affluent journey.

Alhamdulillah.

TABLE OF CONTENT

	Page
CONFIRMATION BY PANEL OF EXAMINERS	ii
AUTHOR'S DECLARATION	iii
ABSTRACT	iv
ACKNOWLEDGEMENT	v
TABLE OF CONTENT	vi
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER ONE INTRODUCTION	1
1.1 Introduction	1
1.2 Research Background	2
1.3 Problem Statements	4
1.4 Research Questions	7
1.5 Objectives	7
1.6 Research Scope	8
1.7 Significance of the Study	9
1.8 Summary	11
CHAPTER TWO LITERATURE REVIEW	12
2.1 Introduction	12
2.2 Phrasing and Prosody	12
2.2.1 Prosodic Structure and Speech Content	13
2.2.2 Boundary Perception	14
2.3 Boundary Labelling	15
2.3.1 Boundary Labels	15
2.3.2 Contour Classification	17
2.4 Boundary Features and Models	19
2.4.1 Duration	20
2.4.2 Phrase Accents	21