Advancing explanatory and tonal dialectometry

Sung, H.W.M.

Advancing Explanatory and Tonal Dialectometry

Published by

LOT
Binnengasthuisstraat 9
1012 ZA Amsterdam
The Netherlands

phone: +31 20 525 2461
e-mail: lot@uva.nl
https://www.lotschool.nl

Cover illustration: A photo of Hong Kong, photographed in December 2019 by Matthew Sung.

ISBN: 978-94-6093-494-0
DOI: https://dx.medra.org/10.48273/LOT0709
NUR: 616

Advancing Explanatory and Tonal Dialectometry

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op vrijdag 13 februari 2026
klokke 11:30 uur

door

Ho Wang Matthew Sung

geboren in 1996
in Hong Kong

Promotor:	Prof. Dr. Yiya Chen
Copromotor:	Dr. Jelena Prokić
Promotiecommissie:	Prof. Dr. Carole Tiberius
	Prof. Dr. Sjef Barbiers
	Prof. Dr. Jack Grieve (University of Birmingham)
	Dr. Wilbert Heeringa (Fryske Akademy)

Contents

Acknowledgments

I owe my gratitude to everyone who has provided me with direct and indirect support throughout my PhD, the preparation of this thesis, and my development as a researcher.

I would first like to thank my brilliant supervisor, Jelena, for her guidance and her non-stopping support, even before my arrival in Leiden (we were in lock down when I started my PhD). I am very grateful for her backing and visions in the field. She is an inspiration to me; she has influenced me a lot in my work and in finding my direction in academia. Throughout my PhD, she has encouraged me to acquire a variety of skills, including programming, multiple sequence alignment, natural language processing and later, even image processing. In retrospect, all of these skills proved to be essential. I would also like to thank my promotor Yiya for her inputs in the tonal part of the thesis. It was not an easy path to start exploring ways to measure tonal distances and her comments to my early ideas and to the modifications of OCO helped a lot in shaping the current thesis. Last but not least, I would like to thank them both for dedicating a lot of time and energy to reading my thesis and providing me with valuable feedback.

I would like to thank my reading committee, Wilbert Heeringa, Jack Grieve, Sjef Barbiers and Carole Tiberius, for reading and evaluating my thesis. Their comments were very helpful in shaping the final draft of my thesis.

A shout out to my proofreaders who have agreed to proofread my thesis without taking any compensations. These generous friends include: Cliodhna Huges, Nina Markl, Tim Espin and Lily Wood. A special mention to Nef and Raoul for proofreading my summaries in Cantonese and

Dutch respectively.

I would also like to mention a number of people who helped me a lot in getting acquainted with dialectometry. Pavel Iosad and Warren Maguire first introduced the concept of dialectometry to me when I was an undergraduate student in Edinburgh. I was very fortunate to have met Wilbert Heeringa at ICLaVE in Leeuwarden in 2019 and he selflessly spent two hours of his dinner time teaching me the basic methods in dialectometry from scratch. I was also fortunate to have the opportunity to meet John Nerbonne in Freiburg at the end of 2019, where he kindly offered to teach me more about using Gabmap for dialectometry. On a different note, I would also like to express my gratitude to Chris Montgomery for teaching me the principles of creating digital maps on QGIS. Without his initial guidance, I would not be able to develop my cartographic skills which I have obtained today. Remco Knooihuizen in Groningen deserves another mention. On the train back to the Netherlands from the Methods conference in Mainz, we had a discussion on how classic dialectometry lacked explainability. This conversation was stuck in my mind until I got the idea of using nPMI to perform feature extraction.

Leiden has become very special to me and one of the reasons is the people I met here. I want to thank my friends from CM, LUCL and LUCDH. Here are a few mentions: Unnur, Laura, Isabella, Priscilla, Aron, Amos, Jiahui, Qi, Marloes, Hans, other houtblazers, the altviool, Yiya's group and colleagues at the LUCDH. I would also like to mention a few friends outside Leiden who have been there for me, even though they are not physically in Leiden, namely Nef, Raoul, Hedwig, Heize, Veronique.

I also want to thank the friends who have helped me to buy books for me in the past few years. Without their help, I won't have access to these valuable resources otherwise.

Last but not least, I want to thank my parents. Since I was a child, they have been giving me pouring love and support. Without them, I would not be able to fulfill my dream of becoming a dialectologist. I am forever grateful for what they have done to support me.

List of Figures

List of Tables