



Universiteit
Leiden
The Netherlands

Exploring tonal variation using dialect tonometry

Sung, H.W.M.; Prokić, J.

Citation

Sung, H. W. M., & Prokić, J. (2024). Exploring tonal variation using dialect tonometry. *Languages*, 9(12). doi:10.3390/languages9120378

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/4209600>

Note: To cite this publication please use the final published version (if applicable).

Article

Exploring Tonal Variation Using Dialect Tonometry

Ho Wang Matthew Sung * and Jelena Prokić *

Leiden University Centre for Linguistics (LUCL)/Leiden University Centre for Digital Humanities (LUCDH),
Leiden University, 2311 EZ Leiden, The Netherlands

* Correspondence: h.w.m.sung@hum.leidenuniv.nl (H.W.M.S.); j.prokic@hum.leidenuniv.nl (J.P.)

Abstract: Most research on dialectometry so far primarily focuses on European languages. Within these studies, analyses on the phonetic level predominantly focus on segments. A lack of studies on languages outside of Europe means that the variation in many lesser-studied languages, including tonal languages, is largely unknown. Tonal languages are languages which pitch is used as an indication in the lexical realisations in (at least some) morphemes, and over half of the world's languages include lexical tones. Despite tones being the inseparable and unneglectable part of the majority of the world's languages, there is only a handful of quantitative dialectometric studies on the dialectal variation in tonal languages. In this paper, we explore the phonetic and phonological variations in Yue, a lesser-studied tonal language spoken by around 80 million people in Southern China. Using a newly devised tone representation (*modified Onset–Contour–Offset*) combined with the Levenshtein distance, we explore the patterns of dialectal variation on the tonal level, as well as to what extent tonal variation correlates with segmental variation. Our results show that tones behave rather differently from segments, and thus, we illustrate that studying lesser-studied and tonal languages can contribute immensely to the study of dialect variation in general.

Keywords: tonal variation; dialect tonometry; dialectometry; linguistic level comparison; computational methods; multidimensional scaling



Citation: Sung, Ho Wang Matthew, and Jelena Prokić. 2024. Exploring Tonal Variation Using Dialect Tonometry. *Languages* 9: 378. <https://doi.org/10.3390/languages9120378>

Academic Editors: Alfred Lameli, Simonetta Montemagni and John Nerbonne

Received: 15 July 2024

Revised: 26 November 2024

Accepted: 28 November 2024

Published: 18 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Quantitative dialectology on phonetics has predominantly been focusing on the segmental level of language variation in recent decades. In addition, the languages which were studied using methods from quantitative dialectology, namely dialectometry, were mainly from Europe, partly due to the development of the field, which centres in Europe. It remains largely unclear whether the findings in European dialectology would also be applicable to regions with rather different historical and sociolinguistic settings.

In recent years, there has been a slight increase in the study of non-European languages in dialectometry, for instance, Yucatan Mayan (Blaha Pfeiler and Skopeteas 2022) and Japanese (Jeszenszky et al. 2019; Heeringa and Inoue 2023). Dialectometric studies have also been conducted on tonal languages, such as those by Yang (2010) and Stanford (2012), but these studies have not addressed tones as a separate linguistic level.

Next to segments (consonants and vowels), tonal languages additionally use tone to distinguish the meaning of one word from another. Hyman (2006, p. 229) provided a definition of tones, which is the use of “pitch... [in] the lexical realisation of at least some morphemes”. It has been estimated that around 60–70% of the languages in the world are tonal (Yip 2002, p. 1). Some relatively well-studied tonal languages include Cantonese, Mandarin, and Yoruba.

Although there has been some work which applied dialectometric techniques on tonal languages, there have been very limited attempts to quantify tonal distances between dialects. Furthermore, these methods were mostly designed for perceptual studies, and most importantly, they were tested with rather small dialect datasets (30 or fewer dialects). This gives rise to the following problem. As Sung et al. (forthcoming) pointed out, due to the

rather low number of dialects in these studies, the tonal inventory of these datasets is not big enough to see how many tones a tone representation can differentiate, which leads to questions of its applicability across languages. After calculating tone distances in four previously described ways in the literature (different tone representations combined with different distance metrics) using a dataset consisting of 104 Yue-Pinghua dialects, they found that most of the existing ways to obtain tone distances cannot even differentiate 50% of the tones in the data. One exception to this generalisation is binary comparison, but it does not yield linguistically coherent distances between tones.

Recently, [Sung et al. \(2024\)](#) proposed a modified version of [Yang and Castro's \(2008\)](#) tone representation, *modified Onset–Contour–Offset* (mOCO hereafter), which has made an improvement in the differentiation of tones in a Yue-Pinghua dialect dataset when combined with the Levenshtein distance. This method of calculating tone distances enables researchers to quantitatively explore tonal variations in (lesser-studied) languages outside of Europe. [Sung et al. \(2024\)](#) applied mOCO on six dialects only to closely study the ways in which tones can vary between dialects.

In the current paper, we build on [Sung et al.'s \(forthcoming\)](#) and previous [Sung et al.'s \(2024\)](#) works on tonal dialectometry by applying mOCO to 104 Yue and Pinghua dialects in order to address two unexplored issues in tonal dialectometry: (1) the existence of geographical dialect boundaries on the tonal level and (2) the comparison of segmental variation to tonal variation, which were not addressed previously to the best of our knowledge. We refer to this type of quantitative dialectology as *dialect tonometry*. This methodology opens the possibility of analysing tonal variation in Southeast Asian and possibly other tonal languages in the world. We hope that our methodology will contribute to the study of dialectal variation in this lesser-studied language area and its position within Sinitic languages.

This paper is structured as follows: Section 2 provides the background of tone languages and how tones are transcribed and presents an overview of the existing works on tonal dialectometry. Section 3 introduces the data and methods used in this study. The results of the analysis on tonal variation are presented in Section 4. Lastly, the discussion and conclusion can be found in Sections 5 and 6, respectively.

2. Tonal Languages

What makes tonal languages different from most European languages is the use of pitch, together with segments, to distinguish the meaning of one word from another. [Hyman's \(2006, p. 229\)](#) definition of tone is the use of “pitch... [in] the lexical realisation of at least some morphemes”. It has been estimated that around 60–70% of the world's languages are tonal ([Yip 2002, p. 1](#)). Some relatively well-studied tonal languages can be found in the Sino-Tibetan language family and the Niger–Congo family, e.g., Cantonese and Yoruba.

The main types of tones are level, falling, and rising tones, as well as complex tones like concave (falling–rising) and convex (rising–falling) tones. A tone is a level tone when the contours are kept flat throughout the whole tone production¹, and a rising tone is when the pitch of the offset is higher than the onset, whereas a falling tone is the opposite. Concave and convex tones are the combinations of rising and falling tones. Languages and dialects may have more than one tone of the same shape. Cantonese, for example, has three level tones, two rising tones, and one falling tone. It has been found that across 737 Sinitic varieties (including over 500 Mandarin dialects), the falling tone is the most common tone (1125 tokens), closely followed by level tones (1086 tokens). Rising (790 tokens) and concave (352 tokens) tones are less common, and convex (80 tokens) tones are the least common ([Cheng 1973](#)).

In this section, the transcription notations of tones (found in dialect surveys) will first be introduced, and then the background and some limits of previous approaches to tonal dialectometry will be provided.

2.1. Transcription Notations in Chinese Dialectology

There are two main types of transcription notations for tones in Chinese dialectology. The first one is a system of notation, which represents the tone contour and pitch levels (phonetically transparent), while the second one is much more (phonetically) opaque.

The transparent notation system aims to show the full tone contour in the transcription. Chao's (1930) *tone letters* is a system for tone transcription which consists of five digits, namely 1, 2, 3, 4, and 5, representing different (possible) contour levels in a tone. In Chao's system, 1 represents the lowest contour level, and 5 represents the highest. When combined (as two-digit or three-digit sequences), they can indicate a change in the contour, which represents the shape of the tone. For example, 53 is a falling tone, whereas 213 is a convex or dipping tone (a falling contour followed by a rising contour). Furthermore, a rising tone (or, inversely, a falling tone) can have three counterparts under this notation, e.g., 35, 335, and 355. These three variants signal canonical, early, and late rises in pitch, respectively.

Chao's (1930) system also has a graphic counterpart, which consists of the contour and a reference height bar on its side. This graphical system is also adapted by the IPA (see the IPA chart, [International Phonetic Association 2005](#)).

Another system found in the Chinese dialectological literature and dialect surveys is more opaque in the sense that these notations do not directly reveal the phonetic realisation of tones. In addition, they are also usually dialect-/language-specific. For instance, it is common in Chinese dialectology to refer to a historical reference system, namely the Middle Chinese (MC hereafter) tone categories, which are *Ping*, *Shang*, *Qu*, and *Ru* (tone category found in checked syllables), and they can be further divided into high and low registers, which are represented by the *Yin* and *Yang* categories, respectively. These categories can be marked as symbols, as shown in Figure 1 (You 2016, pp. 42–44). Set symbols are used to indicate a certain MC tone category depending on where (which corner) it is found in a transcription, as indicated in Figure 1. It should be noted that some dialectologists prefer to use numbers (e.g., '1' represents *Yin Ping*) to represent the MC tone categories instead.

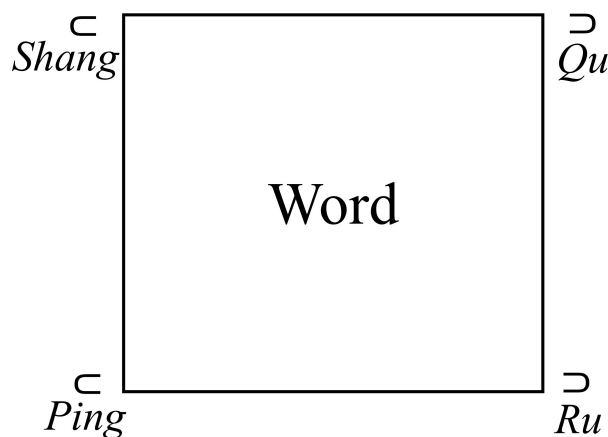


Figure 1. Tone symbols for the Middle Chinese tone categories in the Yin register.

In this paper, tones are referred to in Chao's (1930) notation. This is also the system which most transcriptions in the data use.

2.2. Traditional Approaches to Yue-Pinghua Dialect Tonology

In traditional Chinese dialectology, tones (or tone systems) are often only described in the literature. In Yuan's introduction to major branches of Sinitic languages, the ways tones or tone systems are being described include (1) the number of tones a dialect has (e.g., for Northern/Mandarin dialects; Yuan 2001, p. 37), (2) merger patterns of Middle Chinese sound categories (e.g., the merger of the *Ru* category in Northern/Mandarin di-

alects; Yuan 2001, p. 38), (3) differences between urban–rural and generational variation (e.g., youth speakers do not have the Yang Ru tone in the Nanchang dialect; Yuan 2001, p. 131), (4) correspondences of tone categories between a dialect and the standard (e.g., Meixian vs. Beijing Mandarin; Yuan 2001, p. 163), and (5) correspondences of tone categories between a dialect and another dialect (e.g., Meixian vs. Wuhua; Yuan 2001, p. 164). Seldomly, (6) differences in tone values are discussed (e.g., Zhan 2002, p. 209).

Descriptions of types (1) and (2) are the most common types of descriptions in the literature. A characteristic of these two approaches is that they focus on tone inventories, and they use Middle Chinese tone categories as a reference system. Differences between dialects are revealed when there are differences in the history of these varieties, reflected in various merger (or split) patterns, resulting in different numbers of phonetic tones. These two types of descriptions are also found in Yue-Pinghua tonology. For instance, in Zhan’s (2002, p. 148) description of the Siyi dialects, he wrote that other than the Enping dialect, Siyi dialects have one fewer tone than the Guangzhou dialect.

However, one problem in traditional descriptions is consistency. The level of details decreases for some Sinitic branches. For instance, only the Nanchang dialect was described in Yuan’s (2001) chapter on the Gan branch of Sinitic languages. Here, Nanchang serves as a ‘representative dialect’, hoping to cover features that are present in other dialects of the same language area. The same can be said for Yue. In Zhan’s (2002) description, Yuehai dialects (dialects spoken near Guangzhou) are better described than, e.g., Gao-Lei dialects (dialects spoken in Southwestern Guangdong), where examples from ten dialects of Yuehai are given when discussing similarities and differences, but only two dialects are mentioned, and very brief examples are given.

In terms of differences between a dialect and another dialect/the standard (types 4 and 5), comparisons are often made by listing the MC tone categories of both varieties, and inspections are performed based on how the reflexes map onto each other’s systems. This is achieved qualitatively, and no quantification was carried out with the traditional descriptive approach. Seldomly, comments are made on the differences in tone values, like in Zhan’s (2002, p. 209) description of the tone differences between Guangzhou and Yangjiang dialects.

The main challenge with previous methods for comparing tones among Chinese dialects is the inconsistency in their approaches. Without a systematic description, as well as the level of detail, it is difficult to make comparisons with different sources. In addition, the comparison of tones between different dialects is still restricted to Middle Chinese or the representative dialect and another dialect. It is less common to find comparisons between dialects which are neither representative nor being referred to Middle Chinese (i.e., a synchronic comparison). This may lead to a few problems. Firstly, comparing dialects to the representative dialect alone may not be enough to reveal the degree of similarity/difference between dialects in a given region, making it difficult to obtain an idea of how tones vary between neighbouring (and more distant) dialects. Unless our research interest lies in, e.g., standardisation or the influence of the representative dialect (e.g., in Lau and Leung’s (2011) study on the influence of the Guangzhou dialect on the Ru tone categories), it may not be a suitable method of comparison for the purpose of dialect classification.

There are additional limits to existing traditional approaches. For instance, comparisons are made on the inventory level. However, even if two dialects share the exact same tone inventory, they can still have lexical distributional differences (Sung et al. 2024), which cannot be captured with the traditional approach. Furthermore, it is unclear how tonal comparison can be directly compared with segmental comparison.

To our knowledge, there have not been any dialect classifications (of Yue) which are solely based on tones. Based on the brief overview above, we argue that it is necessary to move beyond qualitative or manual inspections of tone differences, and we should work directly with dialect survey data. Firstly, it will be difficult to ensure consistency in qualitative descriptions of tone differences between a large group of dialects. Secondly, human errors and manual selection of features will be avoided. For these reasons, we argue for the

use of computational distance-based approaches combined with expert analyses. Distance-based approaches have been used for several decades in dialectometry, and there are existing methods which are equipped to deal with a large amount of dialect distances and create visualisations for the analysis. With an appropriate way to calculate tone distances, we can easily integrate the distances into these methodologies in order to gain insights into tonal variation like scholars have with other linguistic levels.

2.3. Previous Approaches to Measuring Tone Distances

Most research on dialectology (descended from the dialect geographical tradition) is conducted on Western languages and other non-tonal languages (including pitch-accent languages like Japanese). Tonal languages and tones have been much less investigated in dialectal research. Furthermore, for tonal languages such as Sinitic languages, most studies are rather descriptive, i.e., they list the tonal inventory of a dialect or show diachronic correspondences with historical tone categories, as in [Zhan’s \(2002\)](#) “Introduction to the Yue dialects in Guangdong”, for example.

The same problem can be found in dialectometry, the measurement of dialect distances using quantitative methods. From [Séguy’s \(1971, 1973\)](#) enterprise of aggregating linguistic features to the application of the Levenshtein distance (e.g., [Heeringa 2004](#)) in calculating phonetic distances, segments have always been the focus. To date, there are not many large-scale dialectometric studies on tones; several studies have looked at a rather limited number of dialects, usually no more than 30 (e.g., [Yang and Castro 2008](#); [Tang 2009](#)). In some dialectometric studies, tones were neglected (e.g., [Wichmann and Ran 2019](#)), and others used binary comparison (e.g., [Stanford 2012](#)).² In addition, there are studies on the correlation between phonetic distance and the perception of tones (e.g., Onset–Contour–Offset (OCO); [Yang and Castro 2008](#)), which do not focus on the application of these measures in dialect classification.

[Sung et al. \(forthcoming\)](#) compared four tone representations in tone distance calculations in the existing literature. These include [Yang and Castro’s \(2008\)](#) OCO, tone-to-string, binary comparison, and the Gandour–Harshman–Tang tone distance ([Tang 2009](#)). An illustration of each method is provided in [Figure 2](#).

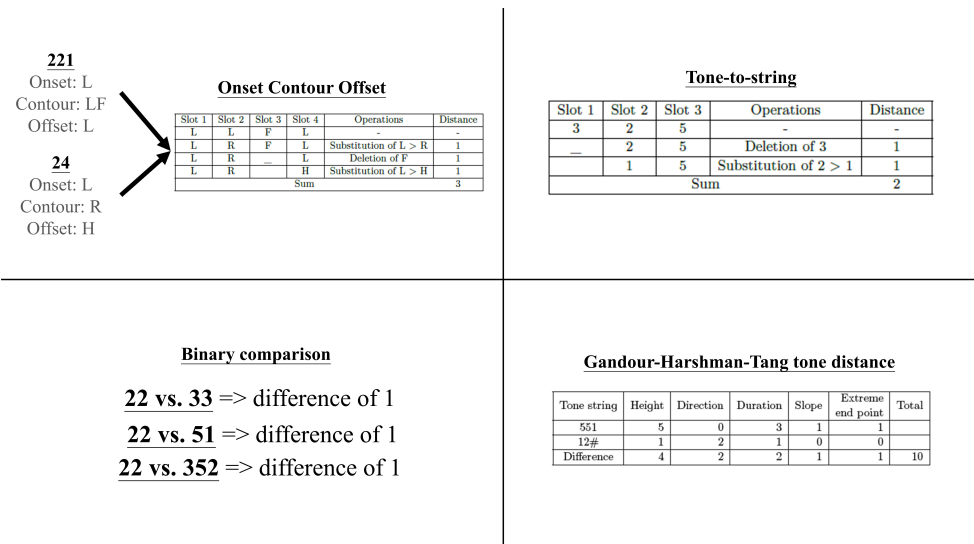


Figure 2. Illustrations of four existing tone distance calculations.

[Sung et al. \(forthcoming\)](#) concluded that none of these ways to obtain tone distances are adequate for measuring tone distances if a dataset consists of a large variety of tones. For instance, most of these methods cannot differentiate all the tones in the dataset of Yue and Pinghua dialects used by [Sung et al. \(forthcoming\)](#), mainly because of the tone representation employed in the methods. This implies that certain tones are counted as the

same, and that tonal variation cannot be examined optimally in a dialectometric analysis. An exception to this is binary comparison. However, binary comparison fails to show any resemblance to cross-linguistic perceptual dimensions in tones. [Gandour and Harshman \(1978\)](#) and [Gandour \(1983\)](#) performed perceptual experiments, and they consistently found that cross-linguistically, the two major perceptual dimensions in tones are average pitch and direction.

A summary of the assessments and their results are shown in Table 1.

Table 1. Overall results of evaluation across 4 tone distance calculations (from [Sung et al. forthcoming](#)).

Method	Overlaps	Perceptual Dimensions	Local Incoherence	ARI (Traditional Classification)	ARI (Segmental Classification)
Binary	None	Uninterpretable	Intermediate	Lowest	Intermediate
Tone-to-string	Intermediate	Uninterpretable	Lowest	Intermediate	Intermediate
OCO	Highest	Resembles GH	Highest	Highest	Highest
GH-T	Intermediate	Identical to GH	Intermediate	Intermediate	Lowest

Based on the results presented in Table 1, Sung et al. argued that OCO and GH-T are preferred representations for tonal dialectometry. However, OCO is perhaps more favourable than GH-T because it is a string-based method and allows for the combination of tonal and segmental analysis. This will not be possible using GH-T since it requires a separate specification of tone features.

Although OCO seems to be a favourable tone representation, Sung et al. also pointed out that OCO’s main problem is that it is unable to distinguish all the tones in the Yue and Pinghua dataset. This is mainly because it only distinguishes three instead of five pitch levels for onsets and offsets. It also does not differentiate short and long phonetic tones. In this paper, the tone representation we use is a modified version of [Yang and Castro’s \(2008\)](#) OCO representation, combined with Levenshtein distance ([Sung et al. 2024](#)). Details of the modifications can be found in Section 3.2.1 below.

Finally, it should be noted that there are other dialectometric studies on suprasegmental level. For instance, [Gooskens and Heeringa \(2006\)](#) included prosody in their distance calculation (by adding stress types to the transcription). Another approach uses F0 contours of sentence intonations extracted from dialect recordings, and then correlations are calculated between the F0 contours of different sentences, and an average is taken between a pair of speakers in order to compute the pairwise ‘similarity’ between the speakers ([Elvira-García et al. 2018](#)). However, these two approaches are not suitable for the current study. Firstly, tones in Norwegian are different from those in Southeast Asian languages. They are called *accentual languages* ([Yip 2002](#)), which usually have one or two contrasting tones, and they are sparsely distributed (or even absent in some words). There is a difference in the way pitch is used between these two types of languages, meaning there is a need for a more sophisticated method for languages with more complex tones. Secondly, the study of intonational dialectometry requires acoustic data and sentence data, which are both absent for the Yue dataset. Hence, the methods discussed in this paper focus on the tone metrics which work on (impressionistic) tonal transcriptions only.

3. Data and Methodology

The data in this study focus on the dialects within the Yue-Pinghua-speaking area. Yue and Pinghua are two sub-branches of the Sinitic branch of the Sino-Tibetan language family, which are spoken in Southern China and the diaspora communities in North America, the UK, Malaysia, etc. ([Wu 2012](#)). Few studies have been conducted on the whole Yue-speaking area with dialectometric methods, especially on the tonal level alone, which is an ideal testing ground for an investigation of spatial variations in tones and comparisons with segments.

3.1. Data

The dataset used in this paper is extracted from varieties of Yue and (Southern) Pinghua (Sung et al. 2024). Yue and Pinghua are mainly spoken in Southern China, namely in the Guangdong and Guangxi provinces, as shown in Figure 3 (based on *Language Atlas of China*, 1st Edition³).

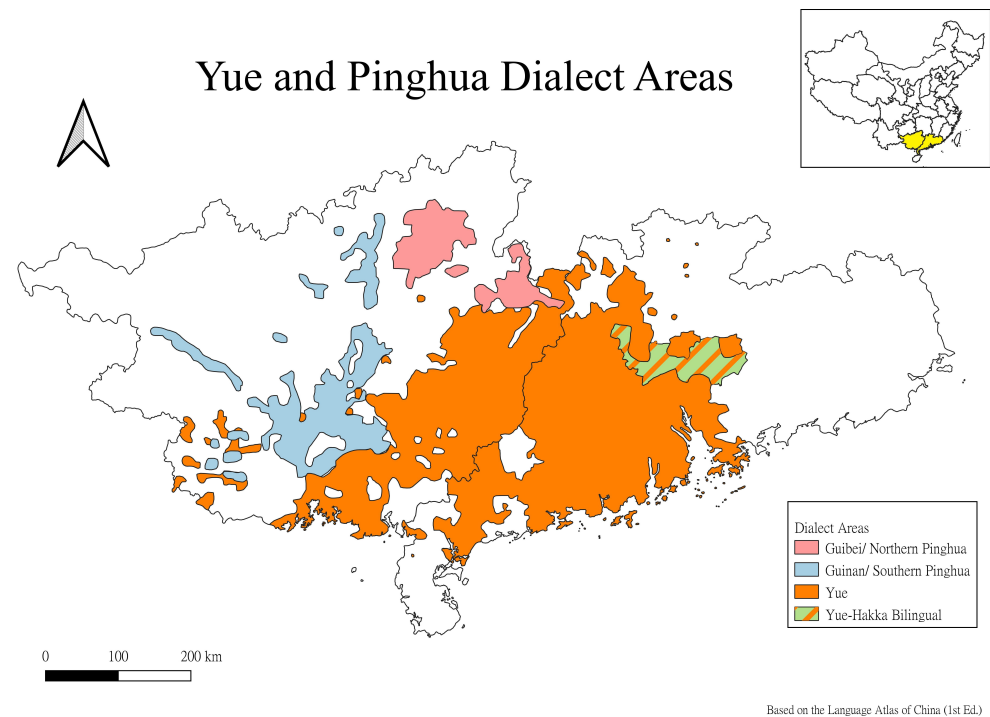


Figure 3. Dialect areas of Yue and Pinghua in Guangdong and Guangxi provinces.⁴

The data consist of impressionistic IPA transcriptions⁵ of monosyllabic words from two types of sources (but both were collected through word list elicitation), namely dialect surveys (Chinese: 方言調查報告) and homonymic syllabaries (Chinese: 同音字滙). The former presents phonetic transcriptions in tabular format (Francis 1983, pp. 105–6), and the latter lists the pronunciation of each character by homonyms (in IPA transcriptions). The dataset only includes monosyllabic words because not all dialect surveys provide polysyllabic words. In addition to the segmental transcriptions, tones are also included in the dataset and in Chao's (1930) tone letters (see Section 2.1).

A subset of the data from seven dialect surveys (Zhan and Cheung 1987, 1994, 1998; Shao 2016; Xie 2007; Chen and Lin 2009; Chen and Weng 2010) was manually extracted and digitised. These dialect surveys were conducted between the 1980s and 2010s, and each of them contains transcriptions of more than 3000 items. Additional data from localities which the surveys did not cover came from mostly homonymic syllabaries⁶ in published sources (Liu 2015; Zhong 2015; Huang 2006; Chen 2009; Yang 2013; Tan 2017; Shi 2009). The localities and the data sources can be found in the map in Figure 4 below.

Around 130 (monosyllabic) words were extracted and digitised from the original sources.⁷ The exact number of words in the database varies as not every dialect contains the same number of transcribed words. The dataset contains 104 Yue and Southern Pinghua varieties.

In total, there are 73 different phonetic tones in the data. A dialect can have 6 to 10 phonetic tones. The tone transcriptions are all impressionistic and not based on any acoustic measurements (the recordings were also not made by the fieldworkers). It is possible that there are transcribers' differences in the data, but since there is no information given about the fieldworkers in the major dialect surveys, it is not possible to assess the effects of the transcribers' differences. However, as described in Section 3.2.1 below, the way we calcu-

late tone distances (mOCO with Levenshtein distance, [Sung et al. 2024](#)) can detect gradual tonal differences (as opposed to binary comparison, where tones are either the same or different). For example, tones 51 and 41, which differ in character ‘L’, would have a distance of 0.2 (see Section 3.2.1) and a distance of 1 when using binary comparison. We hope the gradual difference that mOCO measures can compensate for the potential transcribers’ differences in our data.

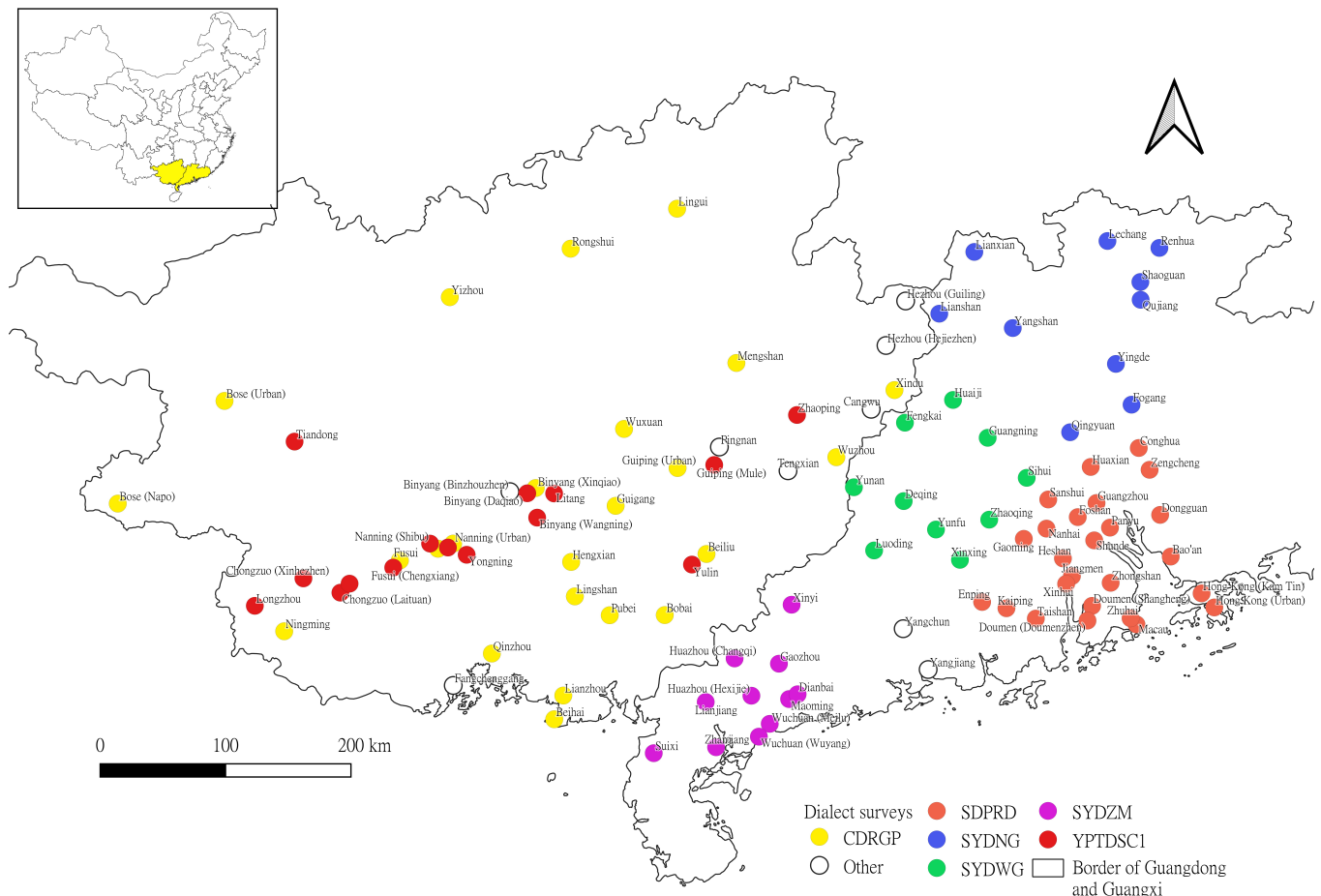


Figure 4. Dialect areas of Yue and Pinghua in Guangdong and Guangxi provinces.

3.2. Dialect Tonometry

In Section 2.2, we pointed out some of the limits in the existing approaches in Chinese and Yue-Pinghua dialect tonology. To address these issues, we opt for a computational approach to tones for the following reasons. In order to move away from comparing the tones between Middle Chinese or Guangzhou and other dialects alone, we propose a method that involves comparing all dialects to one another. This necessitates significant computational resources, as comparing 104 dialects results in 5356 pairwise comparisons, which cannot be carried out manually. This approach allows us to go beyond phonemic inventory comparisons and investigate phonetic and lexical distribution differences. Lastly, methods in dialectometry have been well developed to visualise dialect distances. Calculating dialect distances on the tonal level allows us to explore various questions related to dialect variation in languages, including the patterns of variation across different linguistic levels.

The main goal of this paper is to explore how tones vary geographically within a given area, in this case, the Yue-Pinghua-speaking area in the Guangdong and Guangxi provinces. As pointed out by [Sung et al. \(forthcoming\)](#), since the tone representations in most previous studies are not sufficient in dealing with a large variety of tones often found in bigger dialect datasets, we will use a modified version of OCO ([Sung et al. 2024](#)). This

version of OCO can account for more than 98% of the tones in our data. The measurement of tone distances can be called tonometry, and we decided to label the application of tone distance calculation on the dialects of tonal languages in dialectometry as *dialect tonometry*.

3.2.1. Modified OCO Representation

In the study by Sung et al. (forthcoming), it is shown that OCO is not able to distinguish all tones in the Yue dataset. Sung et al. (2024) focused on this aspect and made some adjustments. Their Modified Onset–Contour–Offset (mOCO) metric still operates with a transformed representation of Chao’s tone letters as string, and the distance is calculated using the Levenshtein distance (Levenshtein 1966; Heeringa 2004). The adjustments are introduced below.

Firstly, the pitch levels are expanded from differentiating three levels (merging 1 and 2 and merging 4 and 5; Yang and Castro 2008) to distinguishing all five levels found in Chao’s (1930) tone letters. The original OCO representation has H vs. M vs. L. Sung et al.’s (2024) modification creates five levels with HH (5), H (4), M (3), L (2), and LL (1). In this representation, the difference between HH and H is one, the same as that between LL and L. The difference between HH/LL and other pitch levels is 2. Additionally, a distinction between short and long tones is made. Differences in tone length are usually found between checked syllables and non-checked syllables, with phonetically short tones represented using a superscript, h. Using mOCO, the difference between short and long tones is 0.5.⁸

Below, we provide an example of how distances between tones 13 and 232 (taken from the tone of 五 ‘five’ in the Guangzhou and Hong Kong (Kam Tin) dialects, respectively) are calculated using mOCO. The mOCO representation of tone 13 is LLRM, and for 232, it is LRFL. To convert LLRM to LRFL using the Levenshtein algorithm, it requires the deletion of L, the insertion of F, and the substitution of M for L. These operations cost a difference of 3 in total. The example is illustrated in Figure 5 below.

	Onset		Contour		Offset	Costs
Tone 13	L	L	R	-	M	
	L	-	R	-	M	Deletion of L
	L	-	R	F	M	Insertion of F
	L	-	R	F	L	Substitution of M for L
Tone 232	L	-	R	F	L	
Costs						3

Figure 5. An example of calculation of tone distance using mOCO representations.

The mOCO representation of the highest and lowest pitch levels maintain the dimensions of ‘direction’ and ‘average pitch’, as shown by Sung et al. (2024, c.f. Table 1). This suggests that the changes added to Yang and Castro’s (2008) OCO can still maintain the perceptual dimensions of tone distances found in Gandour and Harshman’s (1978) perceptual studies.

Multidimensional scaling (MDS) has been used to visualise tone distances found in the Yue data obtained using the mOCO representation and Levenshtein distance. MDS is a technique which can reduce the overwhelming number of distances between all pairs of tones in the data into a more interpretable number of dimensions, usually one or two (a more detailed description can be found in Section 4.1 below). The two-dimensional MDS plot of the distances between the 73 tones in the Yue data can be found in Figure 6. The

modified OCO representation (mOCO hereafter) has shown to improve the distinction of tones in the Yue and Pinghua dataset (Sung et al. 2024). Out of 73 phonetic tones in the dataset, mOCO can now differentiate 72 tones, which is 98.6% of the tones present in our dataset. The only pair of tones mOCO cannot distinguish between is '232' and '242'. Despite the last point, mOCO remains the best tone representation (when combined with the Levenshtein distance) for the exploration of tonal variation in Yue and Pinghua according to Sung et al. (2024).

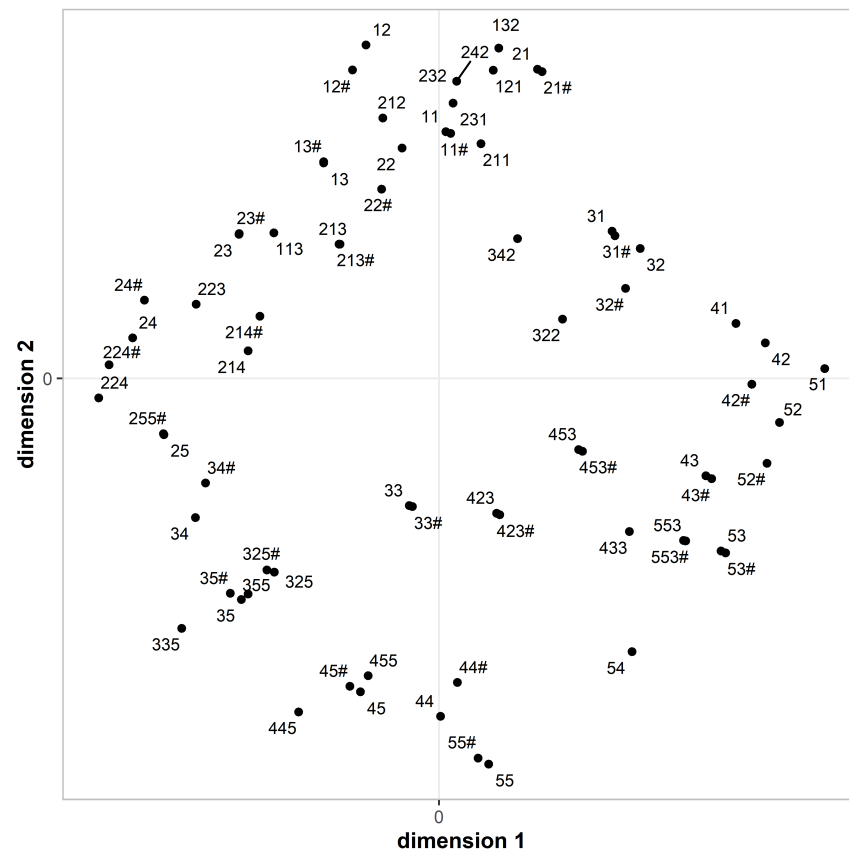


Figure 6. Two-dimensional MDS plot of tone distances using mOCO representation (from Sung et al. 2024). '#' indicate the phonetically short counterpart of a tone.

3.2.2. Aggregate Tone Distances

With an adequate method to calculate individual tone distances, an aggregate tone distance can be calculated between each pair of dialects in the dataset. Obtained aggregate distances can further be analysed using standard methodology from dialectometry to gain more insights into the tonal variation in the Yue and Pinghua dialect area.

One of the important tasks in dialectometry is to identify dialect groups in a specific area. In the so-called aggregate approach to dialect classification, a large sample of dialectal data are compared using methods like *Relative Identity Value* (Goebel 2018) or *Levenshtein distance* (Heeringa 2004) to obtain distances between dialect varieties. This allows the researchers to determine the relationship between dialects in a more objective way and avoid concentrating on a handful of (potentially cherry-picked) features (Nerbonne 2010). These distances are visualised and analysed using various techniques, such as cluster analysis or multidimensional scaling, both of which are independent from geographical information. In addition, the results from the cluster analysis and multidimensional scaling can also be visualised in a map form to gain more insights from a geographical point of view.

In this research, the aggregate distance between every pair of dialects in the dataset is calculated as follows: Firstly, between a pair of dialects, the tone distance between the tones of each lexical item is calculated. Next, the mean of the tone distances of all lexical

items (also known as the aggregate distance) is calculated by summing the tone distances of all lexical items and dividing the sum by the number of lexical items compared. These procedures are repeated for all the pairs of dialects in the dataset.

Lastly, the pairwise distances can be analysed and visualised using a cluster analysis and multidimensional scaling, as well as maps based on the results of these two techniques. The procedures above are illustrated in Figure 7 below.

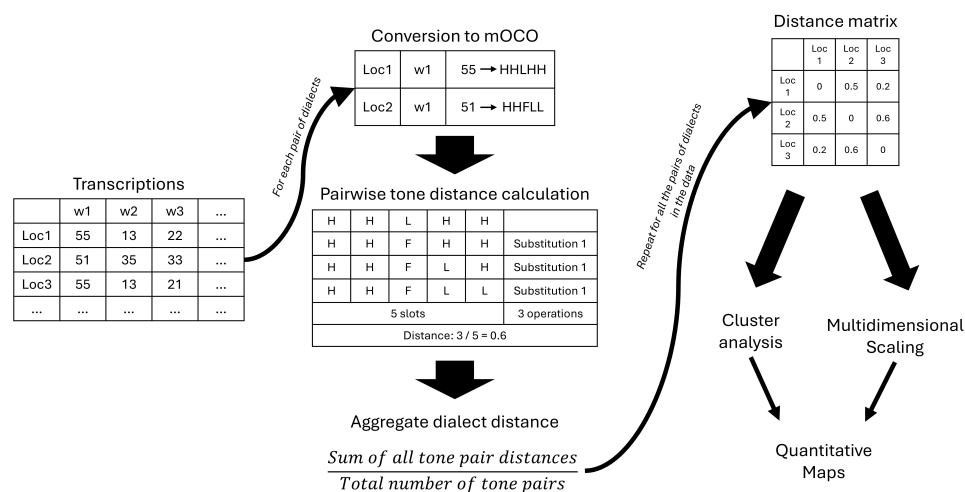


Figure 7. Procedures in dialect tonometry.

4. Tonal Variation Under the Scope of Dialectometry

One of the classical debates in the history of dialectology is whether there are dialect areas (with abrupt boundaries) or a continuum of dialects. The existence of dialect boundaries has been questioned before modern dialect geography (Meyer 1877; Paris 1888), and dialectologists have found evidence for the latter (Chambers and Trudgill 1998; Heeringa and Nerbonne 2001). While comparisons of different linguistic levels, namely lexis, morphology, phonetics, and syntax, have been investigated before on a few European languages (Spruit et al. 2009 for Dutch; Montemagni 2008 for Tuscan dialects; Scherrer and Stoeckle 2016 for Swiss German), tone as a separate linguistic level has not been investigated before in dialectometric studies. This is due to the fact that European languages are not tonal, and dialectometry has not gained full popularity in studies on tone languages.

To the best of our knowledge, the question of whether or not Yue dialects form a dialect continuum has not been addressed in an aggregate analysis. The same applies to tonal variation (i.e., exclusively on tones in a large group of dialects). In this research, we will investigate how tones vary in space (i.e., continuum vs. area) and determine the relationship between tonal vs. segmental variation. First, tonal variation will be examined as a separate system, i.e., without geography involved, following Sung et al.'s (2024) analysis on the Siyi dialects. This analysis provides a perspective on how related or more distantly related dialects vary from each other regardless to where they are spoken. Next, maps are created for the investigation of the relationship between space and tonal variation. Lastly, a comparison between segmental and tonal levels will be carried out in order to build on our understanding of linguistic variation on different linguistic levels.

Since the tonal analysis will be compared with the segmental variation, we will also provide a short analysis of segmental variation so that the comparison is more contextualised. The analysis of segments can be found in Section 4.3 below, following the analyses on tones.

4.1. Tonal Variation in Separate Systems

Do dialects form a continuum on the tonal level in the same way as on the segmental level? It seems that dialects can indeed have gradual differences in terms of tones. Sung

et al. (2024) showed that dialects can vary gradually as individual systems based on the patterns shown in six Siyi dialects. This analysis was conducted using multidimensional scaling so that only the aggregate distances between dialects (as separate linguistic systems) are visualised without accounting for any geographical information (i.e., proximity).

Multidimensional scaling (MDS hereafter) is a dimension reduction method which represents “measurements of similarity (or dissimilarity) among pairs of objects as distances between points” (Borg and Groenen 2005, p. 3). In our case, an MDS plot would represent the dialects as points, and the further the points are from each other, the more different they are. Unlike a cluster analysis, the points on an MDS plot are not partitioned into discrete groups; an MDS plot can show ‘transitional’ dialects, often found in the space of the plot between clusters, which is something a cluster analysis cannot indicate. In addition, no geographical information is added to the procedure or the plot, so the distances projected on the plot are based only on the distance matrix generated in the distance calculation. This technique is useful to determine whether there are distinct groups (clusters) in the data. MDS plots require visual inspections and interpretations of each dimension (see Section 4.4) and the detection of potential clusters.

Figure 8 shows an MDS plot of the aggregate tone distances between Yue and Pinghua dialects (from Sung et al. 2024). There are few small clusters, with the most obvious one being in the lower right corner (circled in red, with Taishan, Enping, and Doumen dialects, which corresponds to traditional Siyi dialects). However, there are no clear tight clusters isolated from the rest of the dialects (other than the Siyi dialects), which suggests a continuum with some loosely defined groups of dialects.

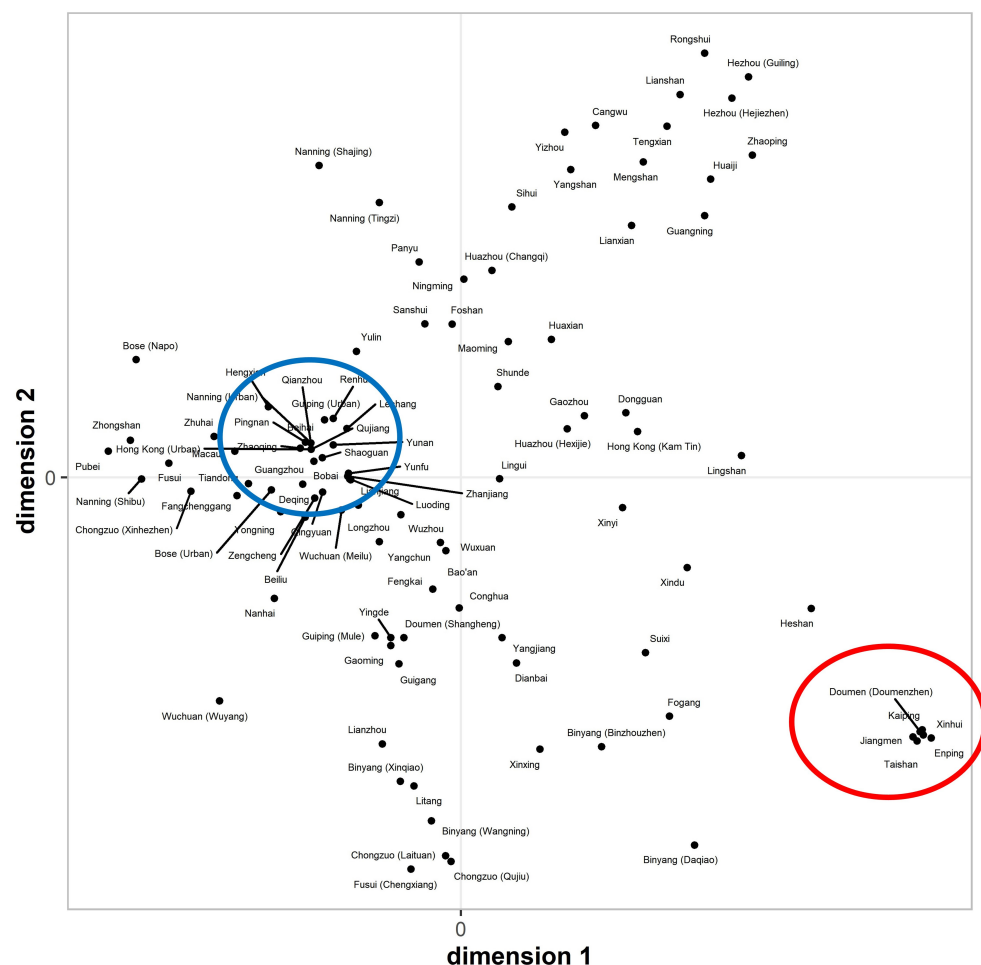


Figure 8. MDS plot of aggregate tone distances between Yue dialects ($r^2 = 0.49$).

Other than the Siyi cluster⁹, another, less prominent, cluster can be found near Guangzhou (circled in blue in Figure 8), which is visualised in Figure 9 below. There are two groups of dialects whose members completely overlap, namely Guangzhou, Hong Kong, Qujiang and Nanning, Beihai, and Qianzhou.¹⁰ This suggests that their tone inventories are completely identical and there are no lexical distribution differences in this dataset. This is a pattern that is different from that in the findings reported in Sung et al.'s (2024) analysis on Siyi dialects, where there are identical inventories but no overlap of dialects on the MDS plot.

We can understand the dialects circled in blue by zooming into the MDS plot near the Guangzhou- and Nanning-type dialects, as well as by looking at how their tone inventories differ. To inspect the tone inventory differences, we use the MC tone categories. MC serves as a historical reference point for comparison instead of using tonemic inventories alone, since tonemes can have the same phonetic contour, while they are found in a completely complimentary sets of words due to their differences in etymological tone categories.

Nanning-type dialects only differ from Guangzhou-type dialects in their reflex of the Yang Shang category. Their tone inventories are listed in Table 2, with Guangzhou and Nanning being representative dialects. Furthermore, there are six dialects which are situated slightly further away from these two groups of dialects. Their tone inventories are also listed in Table 2. The closest dialects to Guangzhou and Nanning are the Pingnan and Bobai dialects. Pingnan is very similar to Guangzhou, except there are a handful of (lexical distributional) differences in the tone correspondences (see Table A1 in the Appendix A). The increasing number of lexical distributional differences as well as differences in the reflex of different tone categories explains why the Pingnan dialect is further away from Guangzhou in the MDS plot than Nanning.

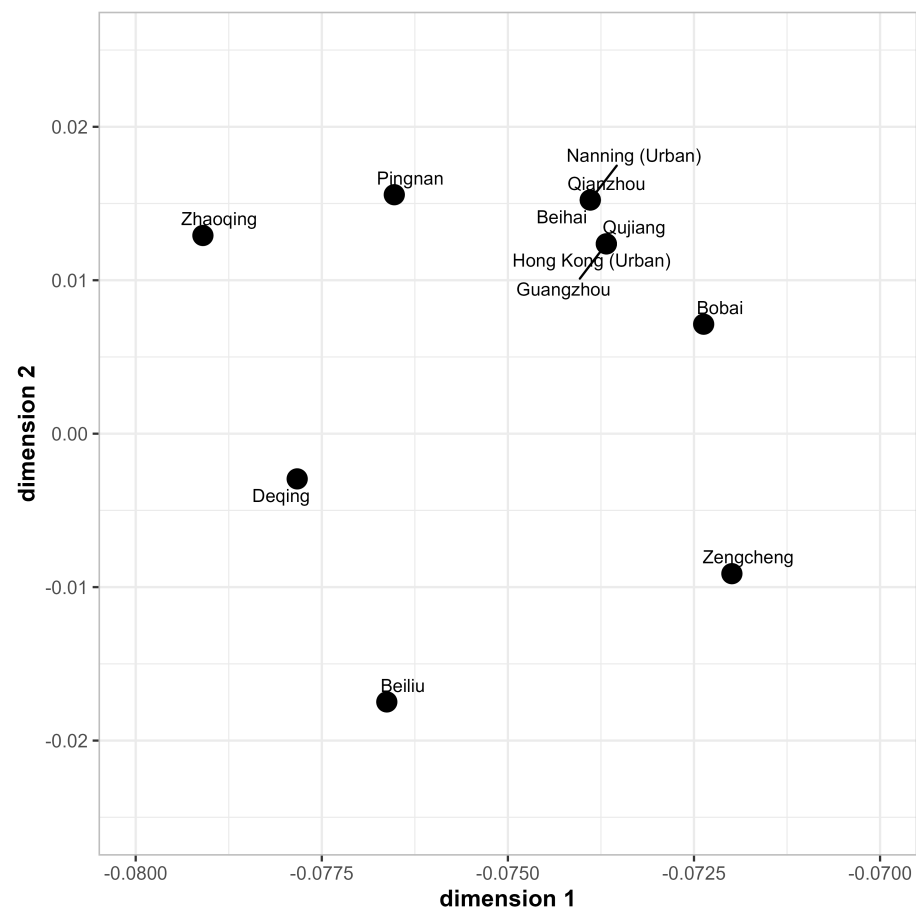


Figure 9. MDS plot of aggregate tone distances between dialects similar to Guangzhou and Nanning.

Next, Zengcheng has a rather similar inventory to Guangzhou and Nanning. Zengcheng has an identical lexical distribution of tones to Guangzhou in the dataset, which means the only difference the MDS plot is reflecting is the difference between Guangzhou 21 and Zengcheng 11 as a reflex of the *Yang Ping* category.¹¹ The remaining three dialects on the MDS plot show different degrees of structural (Zhaoqing does not show contrasts in *Yin Ru1* and *Yin Ru2*), phonetic (e.g., Deqing’s reflex of *Yang Ping* is 342), as well as lexical distributional differences (e.g., in Beiliu), which can then explain why they are even further away from Guangzhou and Nanning.

Finally, the Bobai dialect should be discussed separately for the following reasons. Firstly, a two-dimensional MDS plot does not capture the full variation in tone distances between dialects in our data. This means that some distances cannot be represented in Figure 8, which includes the case for the Bobai dialects in relation to the other dialects. When the third dimension is also accounted for, Bobai differs a lot from the rest of the dialects. On the MDS map in Figure 10, all the dialects mentioned in Figure 8 are green except Bobai, which is in light blue. This is due to the fact that Bobai has a lot of tonal differences (structural, phonetic, and lexical distributions), and we only see that it is located close to Guangzhou due to the limitation of a 2D MDS plot. To show the differences between Bobai and the other dialects, the tonal inventory of the Bobai dialect is also included in Table 2 (as well as the tone correspondences with Guangzhou; see also Table A2 in the Appendix A).

Table 2. Tone inventories of dialects similar to Guangzhou and Nanning dialects (based on Middle Chinese tone categories). Shaded cells indicate the differences discussed in the text. ‘#’ indicate shorted phonetic tones.

Tone Categories	Guangzhou	Nanning	Pingnan	Zengcheng	Zhaoqing	Deqing	Beiliu	Bobai
<i>Yin Ping</i>	55	55	55	55	55	55	55	55
<i>Yang Ping</i>	21	21	21	11	21	342	32	232
<i>Yin Shang</i>	35	35	35	35	35	35	35	33
<i>Yang Shang</i>	13	23	13	13	13	13	23	24
<i>Yin Qu</i>	33	33	53	33	33	53	53	52
<i>Yang Qu</i>	22	22	22	22	52	21	21	21
<i>Yin Ru1</i>	55#	55#	55#	55#	55#	55#	55#	55#
<i>Yin Ru2</i>	33#	33#	33#	33#		33#	33#	33#
<i>Yang Ru1</i>	22#	22#	22#	22#	33#	22#	22#	11#
<i>Yang Ru2</i>								22#

Based on the observation of Figure 9 and the tonal correspondences between these dialects, we can formulate some generalisations. Firstly, the combination of mOCO and the Levenshtein distance can detect not just phonetic but also structural and lexical distributional differences between dialects. Moreover, on an MDS plot, the further away a dialect is from the reference dialect, the more accumulated phonetic, structural, and/or lexical distributional differences there could be. This matches the findings obtained by Sung et al. (2024). This implies that the dialect continuum we see in Figure 8 is also a continuum comprised of accumulated phonetic, structural, and lexical distributional differences.

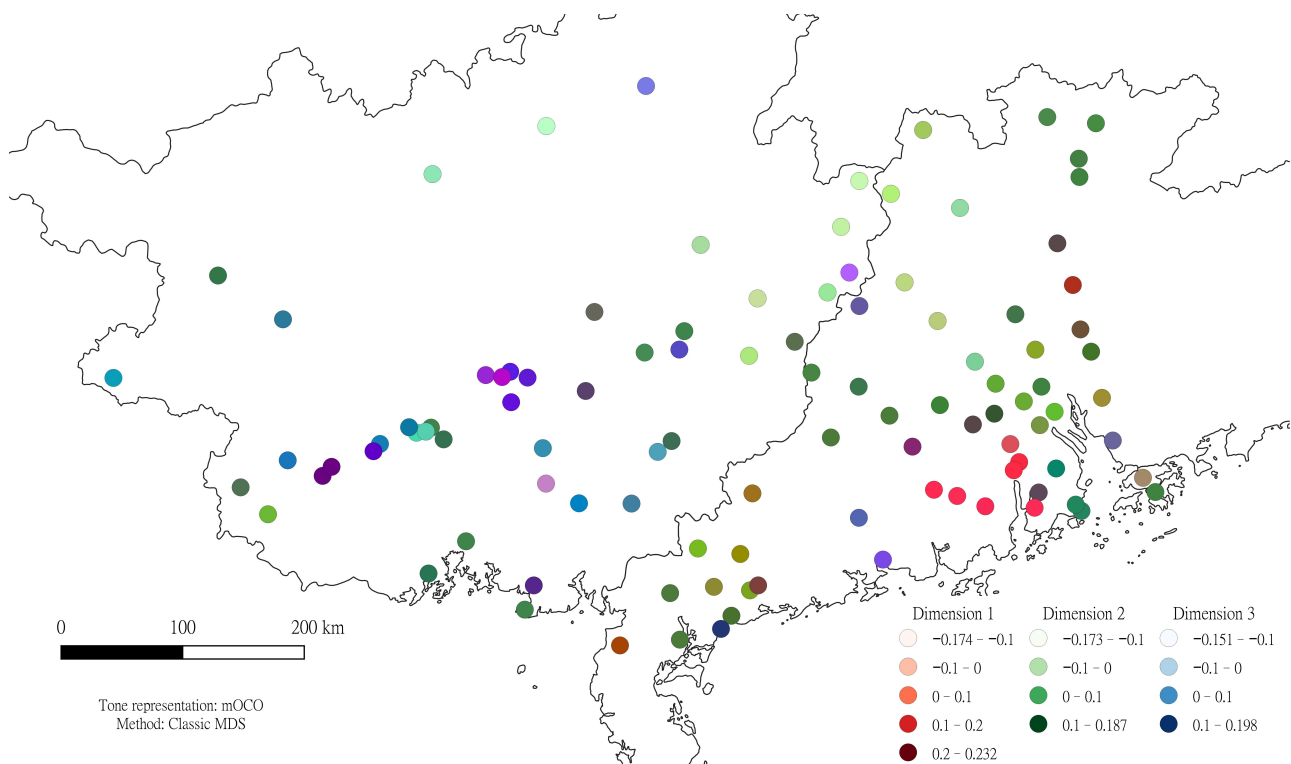


Figure 10. MDS map of tonal variations in Yue dialects ($r^2 = 0.67$).¹²

4.2. A Geographical Continuum?

In Section 4.1, we can see that dialects in our data form a continuum. Since the MDS plot does not include any geographical information, we will project distances on a map to check whether geographical and tonal distances coincide (i.e., whether geographically closer dialects have smaller tonal distances).

4.2.1. MDS Maps

One of the techniques that dialectometrists utilise when visualising the dialect distances across space is the use of MDS maps. The values of the first three MDS dimensions for each dialect are extracted, transformed into RGB colours, and plotted on a map. The MDS map of the tonal distances of Yue can be found below in Figure 10. Dialects that are closer to each other on an MDS plot will have a similar colour on an MDS map; on the contrary, dialects that are distant from each other will have rather different colours. This mapping technique is suitable for a visual inspection of dialect continua and dialect areas since a continuum can be detected through a gradual change in colours, whereas dialect areas can be identified by abrupt colour changes.

A visual inspection of the MDS map (Figure 10) shows several major geographical patterns. One of the areas that stands out is the Siyi dialect area. Unlike the surrounding dialects, Siyi dialects have a scarlet red colour as opposed to the surrounding green and purple-blue circles. This matches the MDS plot in Figure 5 (circled in red) as it is a tight cluster isolated from the rest of the dialects. Next, different shades of green can be found in the eastern inland area, but it extends towards the west to the Zhan-Mao area, i.e., the southwestern part of Guangdong (broken off by some brownish circles) as well as to western Guangxi. Another big cluster of purple circles can be found in central Guangxi, namely around Binyang and Chongzuo. Lastly, a group of dialects clustered together in the north of the Yue area is shown with pale green circles.

4.2.2. Clustering

The tonal distances between the dialects are also analysed using hierarchical clustering. Cluster analysis groups a set of objects in such a way that objects in the same group

(called a cluster) are more similar to each other than to those in other groups (clusters). In this paper, we rely on Ward's hierarchical clustering algorithm (Ward 1963) to extract four clusters and project them on a map (Figure 11).¹³

Four big regions identified using an MDS map resemble the groups identified using a cluster analysis. A cluster map provides a more abrupt border between each cluster (dialect area), and gradual distances between dialects are no longer shown. Other than the four major dialect regions described above, the cluster analysis gives a few geographical outliers, including Lingui, Yangjiang, Yangchun, Dungguan, Xindu, and Fengkai in cluster 3, as well as Fogang and Suixi in cluster 2. Looking more closely, Suixi (the western outlier of Cluster 2) is clustered together with the Siyi dialects. If we look at the tone correspondences between these two dialects, we can see that 60/129 (46.7%) tone pairs are phonetically identical. These two dialects were not often associated together because of their geographical proximity and their segmental differences (as stated in the classifications in the literature). A cluster analysis can highlight distant dialects which share a higher degree of similarity than we expected (based on the Fundamental Dialectological Principle, Nerbonne and Kleiweg 2007). The discovery of these dialects could potentially be useful in uncovering information on contact or historical formation, which is outside the scope of this paper.

What is more striking is that tonal variation does not form a geographical continuum. In Figure 12, we combine MDS and cluster analyses by adding dialect borders based on the cluster analysis to the MDS map presented in Figure 11 (the outliers are ignored). It is not difficult to notice that there are sharp borders between the groups in terms of the differences in colour (which represent dialect distances); red (Siyi) is sharply contrasted with the neighbouring green (Guangfu), light green (Northern) contrasts with the green circles in the south (Guangfu), and the blue-purple circles (Western) contrast with the green. These patterns suggest that tonal variation forms dialect areas rather than a dialect continuum across the Yue-speaking area.

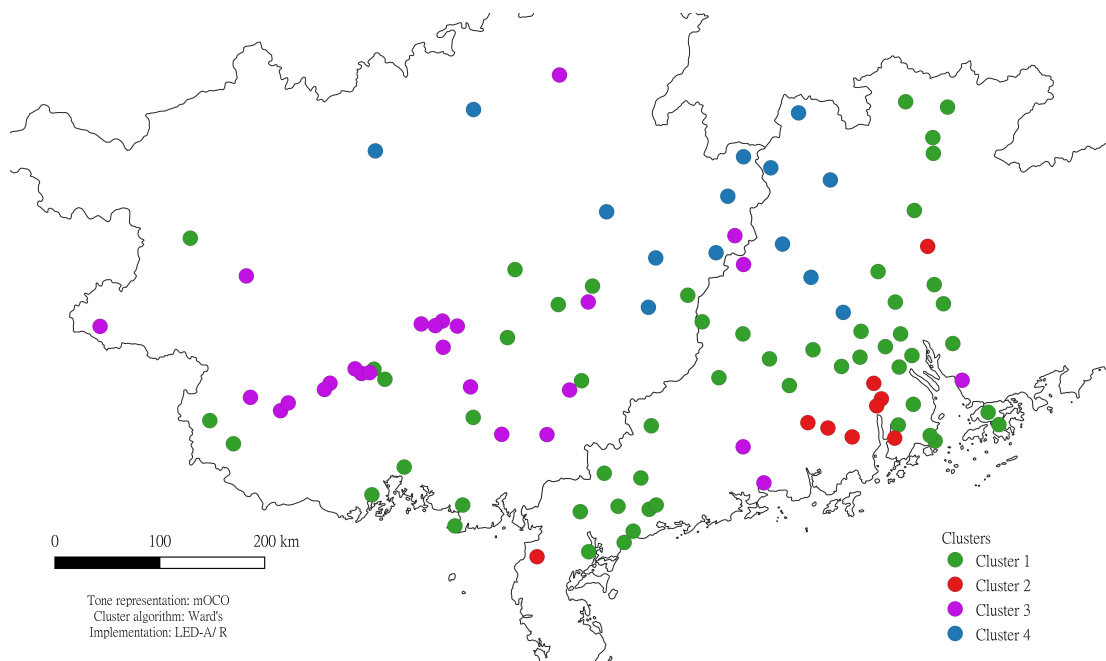


Figure 11. Cluster map of tonal variations in Yue dialects (Ward's method, 4 clusters).

Finally, it should be noted that although dialect areas are found on the tonal level, this does not imply that there are no continua in the data. Within each region, we can still see different shades of green, red, blue-purple, and pale green. This suggests that there are dialect continua within each dialect region.

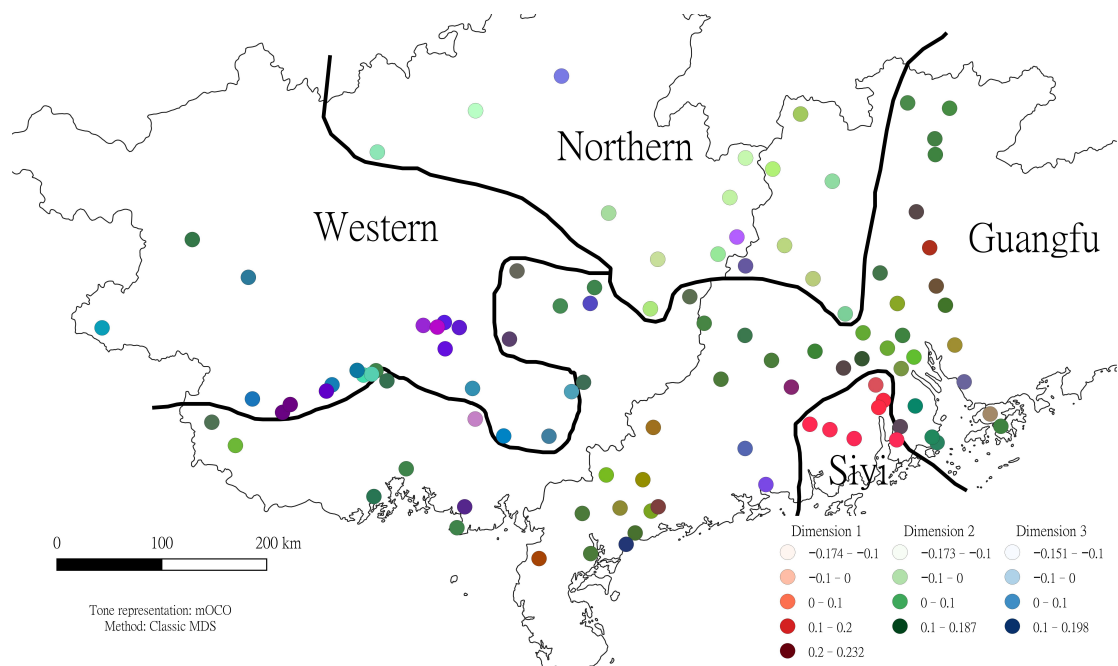


Figure 12. MDS map of tonal variations in Yue dialects with borders based on cluster analysis.

4.3. Segmental Variation in Yue Dialects

The procedures for obtaining and visualising segmental distances of the Yue dialects are largely the same as the procedures shown in Figure 7, except the Levenshtein algorithm is directly applied to impressionistic IPA transcriptions. An example of the calculation of the Levenshtein distance on the segmental level is provided in Figure 13.

Dialects	Onset	Medial	Nucleus	Coda	Operations	Costs
Guangzhou	ts	-	œ	k		
	ts ^h	-	œ	k	Substitution from ts to ts ^h	1
	ts ^h	i	œ	k	Insertion of i	1
	ts ^h	i	ɔ	k	Substitution from œ to ɔ	1
	ts ^h	i	ɔ	-	Deletion of k	1
Bao'an	ts ^h	i	ɔ	-		4

Figure 13. An example of the calculation of the segmental distance using the Levenshtein distance. Red indicates the changed segment.

The segmental distances between Yue dialects are shown on the MDS map in Figure 14. The procedures for creating an MDS map for Yue segmental variation is the same as those used for Figure 10 above, i.e., projecting the dialect distances (in three dimensions) using the RGB colour model. Also, as in Figure 10, the more similar the colours are on this map, the lower the distance between the respective dialects.

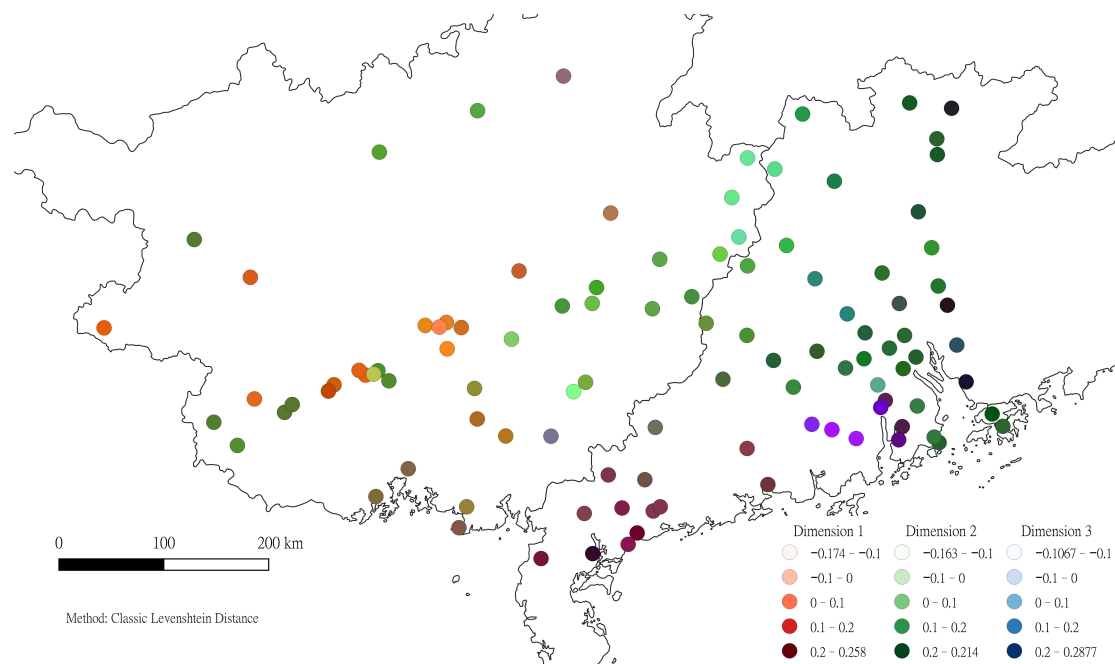


Figure 14. MDS map of segmental variations in Yue dialects ($r^2 = 0.69$).

In Figure 14, we can see a group of dialects in dark green (traditional Guangfu dialects) in the east (Guangdong), which extends all the way to the west (Guangxi). Many of the green dialects in Guangxi are transplanted dialects, formed around 150 years ago during the First Opium War (de Sousa 2020, p. 268). These dialects are mostly distributed along the Xijiang river, resulting in a linear distribution in Figure 10. This is also the reason why there seems to be an abrupt boundary between the green and orange (traditional Guinan Pinghua dialects) dialects in Guangxi. This situation is a bit like an insertion of a newer continuum into an older continuum.

The purple dialects represent the traditional Siyi dialects, which are very distinct from all the dialects in the Yue-speaking area. However, unlike the tonal MDS map in Figure 10, there seems to be transitional dialects between the green (Guangfu) dialects and the Siyi dialects at the eastern border of the Siyi dialects.

Finally, there is another gradual change in colour from green to gray and red near the western coastal area of Guangdong, and the gradual change continues across the border to the eastern coastal area of Guangxi, where there is a transition to brown circles. This gradual change in colour then continues northward towards the orange area, i.e., the traditional Guinan Pinghua varieties. This is also not seen in the tonal MDS map, as there is a distinct boundary between the green and blue/purple dialects for tones (between coastal and inland dialects in Guangxi).

If we ignore the varieties formed around 150 years ago due to migration, we seem to find a continuum that goes from the east to the west of the Yue-speaking area (though there are many gaps in between which could have influenced the analysis, making it look fragmented and less continuum-like). Even between the Siyi and Guangfu dialects, which are very different from each other, we find transitional dialects. This is not seen in tonal variation, as visualised by the two MDS maps.

4.4. Tonal vs. Segmental Variation

Since tones are inseparable from segments on the utterance level, we will compare tonal and segmental distances and their geographical distributions (both taken from Sung et al.'s (2024) dataset). A comparison of different linguistic levels is not new, as it has been carried out in Dutch (Spruit et al. 2009), Swiss German (Scherrer and Stoeckle 2016), and American English (Grieve 2013) between lexis, phonetics, morphology, and/or syntax. However, the relationship between variations on the tonal and segmental levels has not

been explored before. It is intriguing how tones (suprasegments) and segments, which are both part of the phonetic level, pattern.

To compare how much tonal variation correlates with segmental variation, we applied the *Mantel test* (Mantel 1967). The Mantel test is a statistical test which uses permutation to assess the correlation between two distance matrices. The result of the Mantel test also gives a *p*-value which indicates the significance of the correlation based on the permutations. This method was originally proposed to assess the spatial and temporal relationship of a pandemic (Mantel 1967), but it is widely used in other disciplines such as ecology (e.g., Legendre and Legendre 2012). Furthermore, the Mantel test has also been used in dialectometry, namely to investigate the relationship between geographical distances and linguistic distances (Nerbonne and Kleiweg 2007; Huisman et al. 2021).

We performed the Mantel test with our tonal and segmental distance matrices using the R package *Vegan* (Oksanen et al. 2024), which is a package for ecology. To perform the test, we used the *mantel()* function and set the permutation to 9999. The correlation coefficient obtained from the test is $r = 0.52$ with a *p*-value of 0.0001. This indicates that there is a significant positive correlation between tonal variation and segmental variation in the Yue-speaking area, but the strength of the correlation is neither weak nor strong.

To further understand how the two levels of variation are correlated with each other, we performed a correlation analysis between the values from multidimensional scaling for the two distance matrices. This analysis consisted of pairwise comparisons between six groups of values: values from the three dimensions for each linguistic level. Each pairwise comparison returned a Pearson's correlation coefficient and a correlation plot for each pair, which are presented in a correlation matrix, as shown in Figure 15.

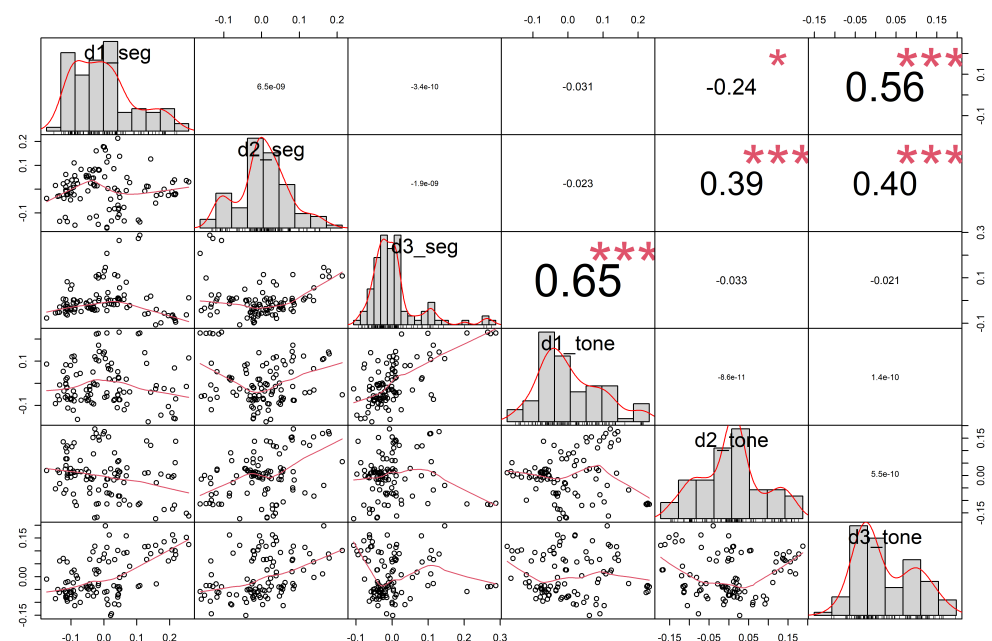


Figure 15. Correlation matrix between each dimension from MDS for tonal and segmental distance matrices. * = $p < 0.05$, *** = $p < 0.001$.

In Figure 15, we can see that the strongest significant correlation can be found between dimension 3 of the segmental matrix and dimension 1 of the tonal matrix, with a correlation coefficient of $r = 0.65$. If we look at the maps in Figure 16, which shows the D3 values for segments (Figure 16a) and D1 values for tones (Figure 16b), we can see that these two dimensions mainly capture the Siyi dialect area. In addition, dialects in the north of the Guangdong–Guangxi border also shows a higher value than most of the dialects in the dataset, although the number of dialects with a higher value differs. These two maps

suggest that both on the tonal and segmental levels, Siyi dialects stand out from the rest of the dialects.

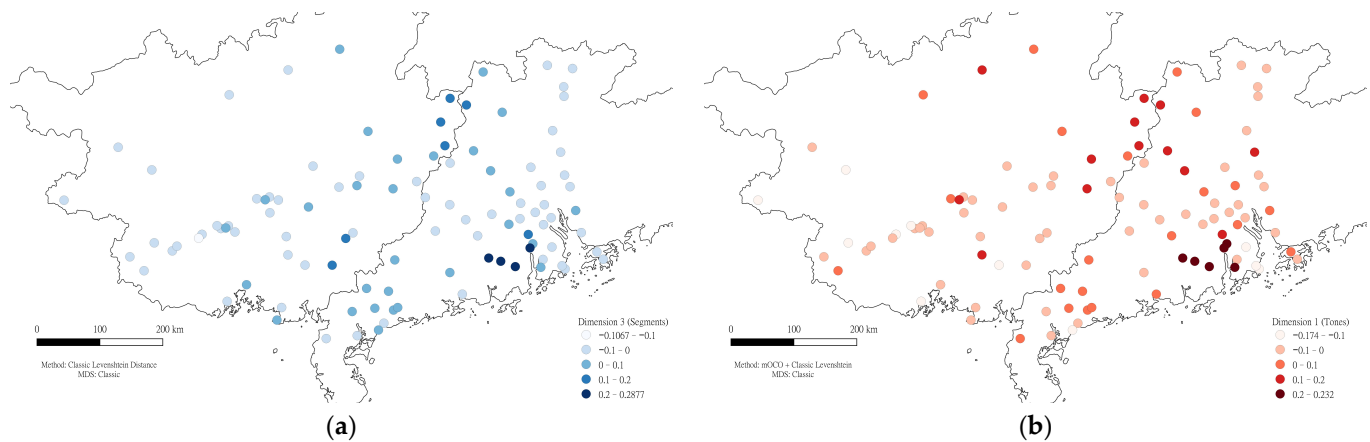


Figure 16. Individual dimension maps for (a) segments (dimension 3; $r^2 = 34.8\%$) and (b) tones (dimension 1; $r^2 = 26.0\%$).

The second strongest correlation found in the correlation matrix lies between dimension 1 of the segments and dimension 3 of tones, with a correlation coefficient of $r = 0.56$. The dialects with the highest dimension values on both levels are found in the Western dialect area (see Figure 17), which correspond to the dialects in the traditional Guinan Pinghua area. Based on the maps, we can also see that there are more Western dialects which differ from the rest of the Yue dialects on the tonal level than the segmental level, but Western dialects are distinctive from the rest of the dialects on both tonal and segmental levels.

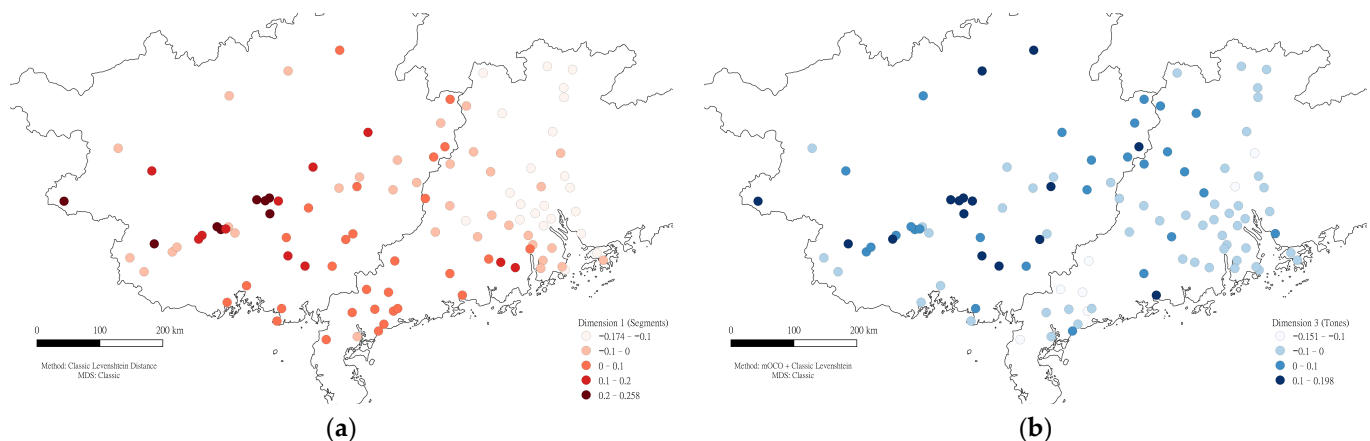


Figure 17. Individual dimension maps for (a) segments (dimension 1; $r^2 = 25.0\%$) and (b) tones (dimension 3; $r^2 = 21.2\%$).

The third strongest pair-wise correlation is between dimension 2 of the segmental level (Figure 18a) and dimension 3 of the tonal level (Figure 17b), and the fourth pair is between dimension 2 of segments (Figure 18a) and dimension 2 of tones (Figure 18b). However, based on the maps shown in Figures 17 and 18, it is unclear which parts of the maps contribute to these correlations.

In general, we can conclude that Siyi and Western dialects are distinct from the rest of the dialects on both linguistic levels. However, for the rest of the dialects, we do not have any immediate evidence that they show a pattern in a similar way, as shown in Figure 18. It should also be noted that in dimension 2 of the tonal analysis, the central northern area stands out from the rest of the dialects, which is not seen in the map for dimension 2 of

the segmental analysis. This might explain why the correlation between the two matrices is only $r = 0.39$ and not any higher.

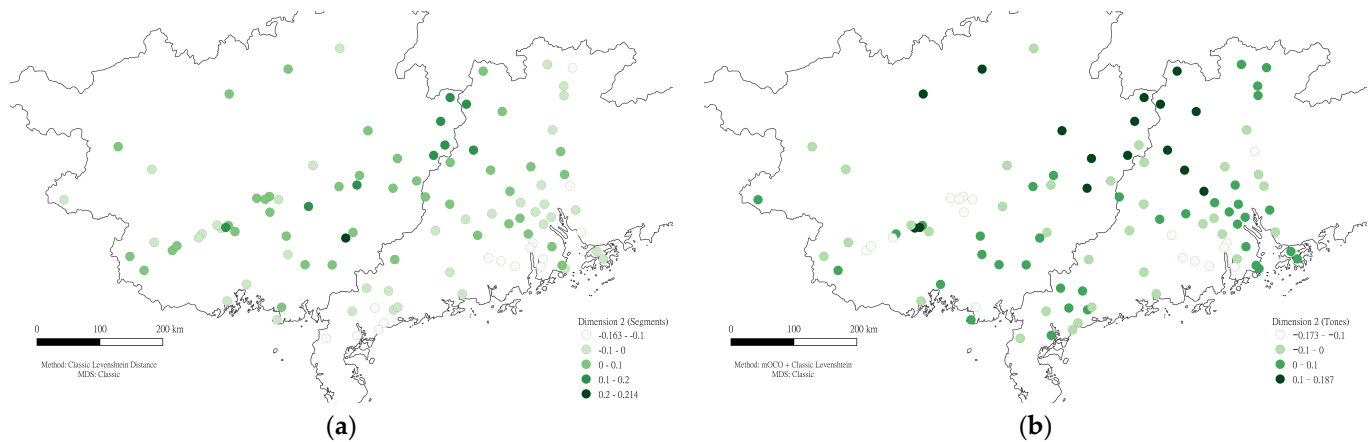


Figure 18. Individual dimension maps for (a) segments (dimension 2; $r^2 = 24.0\%$) and (b) tones (dimension 2; $r^2 = 27.0\%$).

5. Discussion

Based on the MDS map in Figure 10, which reflects the tone distances, we can see that tonal variation is not homogenous. Within a certain dialect area (with the exception of Siyi dialects), there are different shades of green or blue/purple. This is a rather different picture compared to segments (see Figure 13 for the MDS map of segmental variation). This discovery can serve as a reminder for dialectologists that one should not always assume a continuum pattern in dialectal variation.

In Section 4.4, we observed that while two dialect areas, Siyi and Western dialects, correspond to each other, this may not be the case for other areas. It is interesting to see that although segments and tones are both essential parts of the pronunciation of words, they have different degrees of resemblance depending on the dialect area. Guangfu dialects, which have shown an expansion pattern (due to emigration), seem to show a misaligned pattern between the two linguistic levels. One possible explanation could be that tones could change at a different rate from segments in terms of contact-induced changes. The motivation of this hypothesis comes from the fact that in the segmental variation, the MDS map (Figure 14) shows a very clear green continuum going from the east (Guangdong) to the west (Guangxi). Many dialects located in Guangxi, which are spoken near the orange (Western) dialects, are traditionally classified as Yong-Xun dialects, and they were formed around 150 years ago as a result of Guangfu speakers emigrating from the Guangfu region (in the east) following the Western River (de Sousa 2020, p. 268). However, on the tonal MDS map (Figure 10), this emigration pattern is not so clear. Could it be possible that the transplanted varieties were in contact with the local Yue dialects, and segments were somehow transferred quicker than tones? There might be an example of this in the formation of ‘Plastic’ Mandarin in Changsha (Hunan). Plastic Mandarin is a levelled variety between Standard Mandarin and the local Changsha (Xiang) dialect. It has a highly similar segmental inventory to Standard Mandarin (which makes the variety mutually intelligible to Mandarin speakers, while the local Changsha dialect has a distinct phonemic inventory which makes it not intelligible to Mandarin speakers), while the phonetic tones were mainly acquired and adapted from the Changsha dialect (Xu 2022, pp. 151–52). Could this also be the case for some of the Yue dialects in Guangxi? Alternatively, it could also be the other way round, i.e., tones from local dialects influence the transplanted dialects quicker than segments.

Another possibility would be the influence of Zhuang (an indigenous language of the area from the Tai-Kadai family; de Sousa 2020, p. 268). However, the latter explanation cannot explain why there are transitional dialects between Siyi and Guangfu dialects, while on the tonal level, there are very abrupt boundaries between the two groups. Furthermore,

in the Siyi area, it is not known that there is a heavy Zhuang influence of the dialects there (and what does Zhuang have to do with the contact between Guangfu and Siyi, anyway?).

At this point, no conclusions can be made, since this requires more research on segmental vs. tonal changes, as well as a more detailed comparison between Yue and Zhuang dialectal tone systems.

The current paper compares tonal variation vs. segmental variation. The next step would be to explore ways to combine the two levels into a single phonetic level. There are some things to consider: (1) Should we weigh the two levels? If so, how? (2) Should we really combine them simply because they are often grouped as ‘phonetics’, even though they have very different properties, and they show very different behaviours? This is somewhat parallel to arguing whether we should merge morphological variation and syntactic variation together.

Lastly, we would like to make a remark on tone representation. While mOCO is able to differentiate 72 out of 73 tones in the Yue dataset, it is not perfect. It still requires modifications so that differentiation between convex (and concave) tones will be possible, as illustrated in Section 3.2.1.

6. Conclusions

By combining Sung et al.’s (2024) mOCO tone representation with the Levenshtein distance, we explored the patterns in tonal variations in over 100 dialects. We found that dialects can differ in phonetics, the number of phonemes, and lexical distribution, as previously found by Sung et al. (2024). However, unlike previous findings, tones show a pattern which resembles distinct spatial dialect areas rather than a continuum. This does not, however, mean that dialect continua do not exist on the tonal level, since within each dialect area, gradual dialect differences are still found. Lastly, it was found that there is some correlation between tones and segments, although the correlation is not strong. By looking closer at the underlying dimensions between the two matrices, we discovered that the dimensions for the Western dialects and Siyi dialects correlate with each other, but not for the rest of the dialect areas. The mismatch in segmental and tonal variation in the other two dialect groups raises more questions, such as the following: do tones change slower than segments in dialect contact? A future investigation should be conducted on this topic.

We analysed a lesser-studied language, Yue, which is also a tonal language. We hope that there will be more dialectometric studies on other lesser-studied languages, especially tonal languages, in the future. As shown in this study, studying such languages can reveal patterns previously not found in European languages. Given the large number of tonal languages worldwide, more studies would help us better understand dialect variation on the tonal level and contribute to the linguistic theory of tones as well as language change. The same applies to studies of lesser-studied languages, which can provide more insightful knowledge on the nature of variation in different sociolinguistic and socio-historical settings.

Author Contributions: Conceptualization, H.W.M.S.; methodology, H.W.M.S.; software, H.W.M.S.; validation, H.W.M.S. and J.P.; formal analysis, H.W.M.S.; investigation, H.W.M.S.; resources, H.W.M.S.; data curation, H.W.M.S.; writing—original draft preparation, H.W.M.S.; writing—review and editing, H.W.M.S. and J.P.; visualization, H.W.M.S.; supervision, J.P.; project administration, J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Yue-Pinghua dataset is available here: <https://osf.io/m9g2a/>.

Acknowledgments: We thank Yiya Chen (Leiden University) for her comments for the mOCO representation.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Correspondence table of tones between Guangzhou (left) and Pingnan (right) dialects (irregular correspondences in gray).

Correspondences	Tokens
13:13	16
21:13	1
21:21	19
22:21	1
22:22	9
22#:22#	12
22#:255#	1
33:53	7
33#:33#	7
33#:55#	1
35:35	18
35:55	1

Table A2. Correspondence table of tones between Guangzhou (left) and Bobai (right) dialects (irregular correspondences in gray).

Correspondences	Tokens
13:24	15
13:33	1
21:232	21
22:21	10
22#:11#	2
22#:22#	11
33:52	7
33#:33#	7
33#:55#	1
35:33	18
35:55	1
55:55	28
55#:55#	8

Notes

- ¹ In reality, some level tones also have a slight falling or rising F0 trajectory, but this paper focuses on impressionistic transcription data, which do not always capture such fine trajectories.
- ² The tone of a word is indicated by the tone category, which represents a phonetic tone value. However, the author only compares whether or not the tone category is the same (binary comparison). The tone is considered the same if both tone categories in a pairwise comparison are identical, even if their phonetic values are different (Stanford 2012, p. 261). If the tone categories are different, even if the tone values are the same, they are still considered different.
- ³ The first edition (Li and Wurm 1987) is used here because the second edition (Chinese Academia of Social Science/CASS 2012) is based on county boundaries instead of the approximation of the dialect boundaries, which gives more unprecise delimitations of the dialect areas compared to this map.
- ⁴ All maps in this paper were produced with QGIS (QGIS Development Team 2023).
- ⁵ A modified version of the IPA is used in Chinese dialectology, with a number of non-standard characters, such as the apical vowel [ɿ], which roughly represents [z]; [ʰ] for aspiration; and [A] and [E] for roughly [a_] and [e_]/[ɛ_] (between [e] and [ɛ]), respectively (Handel 2015).
- ⁶ A list of (monosyllabic) words organised by their pronunciation; words are grouped by their rhymes, and then their onset, and finally, by their tone.
- ⁷ Due to the time constraint for manual data extraction and digitisation, only 130 items were digitised.
- ⁸ Superscripted characters are counted as having a difference of 0.5 in the Levenshtein algorithm implemented on Gabmap (up until the first half of 2024, when the authors rechecked, Nerbonne et al. 2011; Leinonen et al. 2016) if the last character of the offset of both tones (but not the length) are identical. This implies that the tone length is only differentiated if the final characters

of the offset in the mOCO representation are identical. Please note that *LED-A.org* does not have the same implementation of the superscript *h*.

The Siyi dialects were already analysed by Sung et al. (2024). Please refer to their paper for the analysis of these six dialects.

It should be noted that these dialects are not neighbouring dialects.

The number 21 is labelled as a falling tone, and 11 is labelled as a level tone in mOCO, which is the main cause of the difference.

This map and the following maps are zoomed into the area of interest. The blank spaces, which include mostly Guangdong but also part of Guangxi and consist of linguistic areas which are not Yue-speaking, are omitted for a better view of the variation in the studied area.

A cluster analysis was performed in R (R Core Team 2024).

References

- Blahe Pfeiler, Barbara, and Stavros Skopeteas. 2022. Sources of convergence in indigenous languages: Lexical variation in Yucatec Maya. *PLoS ONE* 17: e0268448. [CrossRef] [PubMed]
- Borg, Ingwer, and Patrick J. F. Groenen. 2005. *Modern Multidimensional Scaling: Theory and Applications*. New York: Springer Science and Business Media.
- Chambers, Jack K., and Peter Trudgill. 1998. *Dialectology*, 2nd ed. Cambridge: Cambridge University Press.
- Chao, Yuen-Ren. 1930. A system of tone letters. *Le Maître Phonétique* 8: 24–27. Available online: <http://www.jstor.org/stable/44704341> (accessed on 6 February 2021).
- Chen, Hailun, and Yi Lin. 2009. *Yue yu ping hua tu hua fang yin zi hui di 1 pian: Guangxi yue yu, Guinan ping hua bu fen* 粵語平話土話方音字彙第1編: 廣西粵語、桂南平話部分 [The Lexicon of Yue, Pinghua and Tuhua Volume 1: Guangxi Yue and Guinan Pinghua]. Shanghai: Shanghai Education Publishing.
- Chen, Xiaoyan. 2009. 廣西賀州八步(桂嶺)本地話音系 [The Phonology of Hezhou Babu (Guiling) Local Vernacular in Guangxi]. *方言 [Dialect]* 1: 53–71.
- Chen, Xiaojin, and Zewen Weng. 2010. 粵語西翼考察—廣西貴港粵語之個案研究 [An Investigation in the West of the Yue-speaking Area—A Case-Study on Guigang Yue in Guangxi]. Guangzhou: Jinan University Press.
- Cheng, Chin-Cuan. 1973. A quantitative study of Chinese tones. *Journal of Chinese Linguistics* 1: 93–110. Available online: <http://www.jstor.org/stable/23749778> (accessed on 18 July 2018).
- Chinese Academia of Social Science/CASS. 2012. 中國語言地圖集 [Language Atlas of China], 2nd ed. Beijing: Commercial Press.
- de Sousa, Hilario. 2020. On Pinghua and Yue: Some Historical and Linguistic Perspectives. *Crossroads* 19: 257–95. [CrossRef]
- Elvira-García, Wendy, Simone Balocco, Paolo Roseano, and Ana M. Fernandez-Planas. 2018. ProDis: A dialectometric tool for acoustic prosodic data. *Speech Communication* 97: 9–18. [CrossRef]
- Francis, W. Nelson. 1983. *Dialectology: An Introduction*. Boston: Addison-Wesley Longman Limited.
- Gandour, Jackson. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics* 11: 149–75. [CrossRef]
- Gandour, Jackson, and Richard A. Harshman. 1978. Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech* 21: 1–33. [CrossRef]
- Goebel, Hans. 2018. Dialectometry. In *The Handbook of Dialectology*. Edited by Charles Boberg, John Nerbonne and Dominic Watt. Hoboken: John Wiley & Sons.
- Gooskens, Charlotte, and Wilbert Heeringa. 2006. The relative contribution of pronunciation, lexical, and prosodic differences to the perceived distances between Norwegian dialects. *Literary and Linguistic Computing* 21: 477–92. [CrossRef]
- Grieve, Jack. 2013. A statistical comparison of regional phonetic and lexical variation in American English. *Literary and Linguistic Computing* 28: 82–107. [CrossRef]
- Handel, Zev. 2015. Non-IPA Symbols in IPA Transcriptions in China. In *Encyclopedia of Chinese Language and Linguistics*. Edited by Rint Sybesma. Leiden: Brill. [CrossRef]
- Heeringa, Wilbert. 2004. Measuring Dialect Pronunciation Using Levenshtein Distance. Ph.D. dissertation, University of Groningen, Groningen, The Netherlands.
- Heeringa, Wilbert, and Fumio Inoue. 2023. Exploring the Japanese Dialect Geography Dialectometrically: Division and Continuity. *Studies in Geolinguistics* 3: 1–44.
- Heeringa, Wilbert, and John Nerbonne. 2001. Dialect areas and dialect continua. *Language Variation and Change* 13: 375–400. [CrossRef]
- Huang, Qun. 2006. 賀州市賀街本地話同音字匯 [The Homonymic Syllabary of the Hezhou Local Vernacular in Hezhou]. *Journal of Guilin Normal College* 20: 6–13.
- Huisman, John L. A., Karlien Franco, and Roeland van Hout. 2021. Linking Linguistic and Geographic Distance in Four Semantic Domains: Computational Geo-Analyses of Internal and External Factors in a Dialect Continuum. In *Computational Sociolinguistics. Frontiers*. Edited by Jack Grieve, Dirk Hovy, David Jurgens, Tyler S. Kendall, Dong Nguyen, James N. Stanford and Meghan Sumner. Available online: <https://www.frontiersin.org/about/contact> (accessed on 18 July 2018).
- Hyman, Larry M. 2006. Word-prosodic typology. *Phonology* 23: 225–57. [CrossRef]
- International Phonetic Association. 2005. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

- Jeszszsky, Péter, Yoshinobu Hikosaka, Satoshi Imamura, and Keiji Yano. 2019. Japanese lexical variation explained by spatial contact patterns. *ISPRS International Journal of Geo-Information* 8: 400. [CrossRef]
- Lau, Chun-Fat, and Wai-Mun Leung. 2011. 珠江三角洲方言陰入聲受廣州話影響的演變情況 [The Influence of Cantonese on the Yinru(陰入) Tone of the Pearl River Delta Dialects]. *Studies in Language and Linguistics* 31: 92–97.
- Legendre, Pierre, and Louis Legendre. 2012. *Numerical Ecology*, 3rd ed. Amsterdam: Elsevier.
- Leinonen, Therese, Çağrı Çöltekin, and John Nerbonne. 2016. Using gabmap. *Lingua* 178: 71–83. [CrossRef]
- Levenshtein, Vladimir I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady* 10: 707–10.
- Li, Rong, and Stephen A. Wurm. 1987. 中國語言地圖集 [Language Atlas of China]. Hong Kong: Hong Kong Longman Group.
- Liu, Caifeng. 2015. 廣東兩陽粵語語音研究. [A Phonological Study on Cantonese in Liangyang Area of Guangdong]. Ph.D. Dissertation, Jinan University, Jinan, China.
- Mantel, Nathan. 1967. The detection of disease clustering and a generalised regression approach. *Cancer Research* 27: 209–20. [PubMed]
- Meyer, Paul. 1877. Archives des missions scientifiques et littéraires. 3e série, t. III, 2e livraison. *Romania* 6: 630–33. Available online: https://www.persee.fr/doc/roma_0035-8029_1877_num_6_24_6844 (accessed on 27 September 2021).
- Montemagni, Simonetta. 2008. The space of Tuscan dialectal variation. A correlation study. *International Journal of Humanities and Arts Computing* 2: 135–52. [CrossRef]
- Nerbonne, John. 2010. Mapping aggregate variation. *An International Handbook of Linguistic Variation* 2: 476–95.
- Nerbonne, John, and Peter Kleiweg. 2007. Toward a dialectological yardstick. *Journal of Quantitative Linguistics* 14: 148–66. [CrossRef]
- Nerbonne, John, Rinke Colen, Charlotte Gooskens, Peter Kleiweg, and Therese Leinonen. 2011. Gabmap—a web application for dialectology. *Dialectologia: Revista Electrónica* II: 65–89. Available online: http://www.edicions.ub.edu/revistes/ejecuta_descarga.asp?codigo=717 (accessed on 18 July 2018).
- Oksanen, Jari F., Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Sölymos, M. Henry, H. Stevens, and et al. 2024. Vegan: Community Ecology Package. Software. Available online: <http://CRAN.R-project.org/package=vegan> (accessed on 18 July 2018).
- Paris, Gaston. 1888. Les parlers de France. *Revue des Patois Gallo-Romans* 2: 161–75.
- QGIS Development Team. 2023. QGIS Geographic Information System. Open Source-Geospatial Foundation Project. Available online: <http://qgis.osgeo.org> (accessed on 18 July 2018).
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 18 July 2018).
- Scherrer, Yves, and Philipp Stoeckle. 2016. A quantitative approach to Swiss German–Dialectometric analyses and comparisons of linguistic levels. *Dialectologia et Geolinguistica* 24: 92–125. [CrossRef]
- Séguy, Jean. 1971. La relation entre la distance spatiale et la distance lexicale. *Revue de Linguistique Romane* 35: 335–57. [CrossRef]
- Séguy, Jean. 1973. La dialectométrie dans l'atlas linguistique de la Gascogne. *Revue de Linguistique Romane* 37: 1–24. [CrossRef]
- Shao, Huijun. 2016. 粵西湛茂地區粵語語音研究 [The Phonological Study of the Yue Dialects spoken in the Zhan-Mao area in Western Guangdong]. Guangzhou: Sun Yat-Sen University Press.
- Shi, Rimei. 2009. 廣西防城區粵語音系 [The Phonology of the Fangcheng Yue dialect in Guangxi]. 百色學院學報 [Journal of Baise University] 22: 106–16.
- Spruit, Marco R., Wilbert Heeringa, and John Nerbonne. 2009. Associations among linguistic levels. *Lingua* 119: 1624–42. [CrossRef]
- Stanford, James N. 2012. One size fits all? Dialectometry in a small clan-based indigenous society. *Language Variation and Change* 24: 247–78. [CrossRef]
- Sung, Ho Wang Matthew, Jelena Prokić, and Yiya Chen. 2024. A New Dataset for Tonal and Segmental Dialectometry from the Yue- and Pinghua-Speaking Area. In *Proceedings of the 6th Workshop on Research in Computational Linguistic Typology and Multilingual NLP*. pp. 25–36. Available online: <https://aclanthology.org/2024.sigtyp-1.3.pdf> (accessed on 18 July 2018).
- Sung, Ho Wang Matthew, Jelena Prokić, and Yiya Chen. Forthcoming. Applying the state-of-the-art tonal distance metrics to a large dialectal dataset. In *(Dia)lects in the 21st Century: Selected Papers from Methods in Dialectology XVII (Mainz, 2022) (Language Variation)*. Edited by Wagner Susanne and Stange-Hundsdoerfer Ulrike. Berlin: Language Science Press.
- Tan, Yuanxiong. 2017. 廣西賓陽縣(賓州鎮)本地話音系 [The Phonology of Binzhouzhen in the Binyang County in Guangxi]. 梧州學院學報 [Journal of Wuzhou University] 27: 58–71.
- Tang, Chaoju. 2009. Mutual Intelligibility of Chinese Dialects: An Experimental Approach. Ph.D. dissertation, Leiden University, Leiden, The Netherlands.
- Ward, Joe H., Jr. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58: 236–44. [CrossRef]
- Wichmann, S., and Q. Ran. 2019. ASJP模式的漢語方言計算分析—以65個漢語方言語檔為例 [A Phylogenetic Study on 65 Chinese Doculects: With ASJP Tools]. 現代語文 [Modern Languages] 5: 4–13.
- Wu, Wei. 2012. 粵語 [Yue]. In *Chinese Academia of Social Science/CASS. 中國語言地圖集 [Language Atlas of China]*, 2nd ed. Beijing: Commercial Press.
- Xie, Jianyou. 2007. *Guangxi han yu fang yan yan jiu* 廣西漢語方言研究 [Studies on the Chinese dialects in Guangxi]. Nanning: Guangxi People's Publishing.
- Xu, Chenzi. 2022. Investigating the Tonal System of Plastic Mandarin: A Cross-Varietal Comparison. Ph.D. thesis, University of Oxford, Oxford, UK.

- Yang, Cathryn. 2010. Lalo Regional Varieties: Phylogeny, Dialectometry, and Sociolinguistics. Ph.D. dissertation, La Trobe University, Bundoora, Australia.
- Yang, Cathryn, and Andy Castro. 2008. Representing tone in Levenshtein distance. *International Journal of Humanities and Arts Computing* 2: 205–19. [[CrossRef](#)]
- Yang, Shiwen. 2013. 廣西藤縣濛江方言音系 [The Phonology of the Tengxian Mengjiang Dialect in Guangxi]. *方言 [Dialect]* 1: 71–85.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.
- You, Rujie. 2016. 漢語方言學教程 [A Course in Chinese Dialectology]. Shanghai: Shanghai Educational Publishing House.
- Yuan, Jiahua. 2001. 漢語方言概要 [An Outline of Chinese Dialects]. Beijing: Yu Wen Publishing House.
- Zhan, Bohui, ed. 2002. 廣東粵方言概要 [Introduction to the Yue dialects in Guangdong]. Jinan: 暨南大學出版社.
- Zhan, Bohui, and Yat-Shing Cheung. 1987. *A Survey of Dialects in the Pearl River Delta, Vol. 1, Comparative Morpheme-Syllabary*. Guangzhou: People's Publishing House of Guangdong.
- Zhan, Bohui, and Yat-Shing Cheung. 1994. *A Survey of Yue Dialects in North Guangdong*. Jinan: Jinan University Press.
- Zhan, Bohui, and Yat-Shing Cheung. 1998. *A Survey of Yue Dialects in West Guangdong*. Jinan: Jinan University Press.
- Zhong, Ziqiang. 2015. 廣西蒼梧本地話音系 [The Phonology of Cangwu Local Vernacular in Guangxi]. *方言 [Dialect]* 2: 177–92.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.