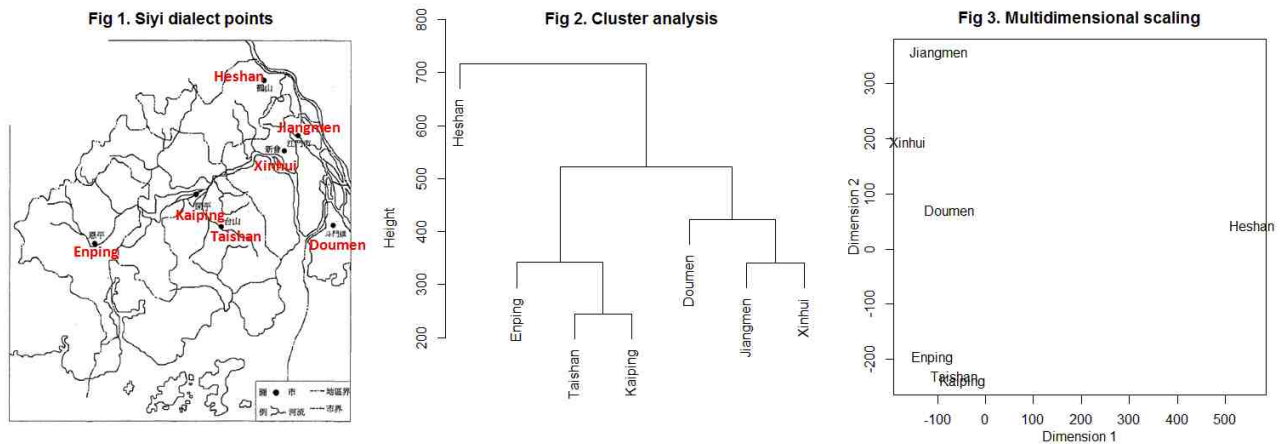


Quantifying the lexical affinity among Siyi Yue dialects

Yutian Tan & Tsz-Him Tsui
The Ohio State University

Introduction: Previous research mainly focused on the differences between Standard Cantonese and Siyi Yue dialects, a subgroup of Yue spoken in the southwestern area in the Pearl River Delta in Guangdong Province of China. Only a few studies (Gan 2002, Tong 2004) have tackled the internal variation within Siyi dialects, while none of them used quantitative approaches. This current study quantifies lexical differences between the Siyi dialects, and presents a synchronic classification of Siyi. The current classification conforms with the geographic distances and contact situations between the dialect points.

Methods: The current data contains 1340 lexical items documented in Zhan & Cheung (1988). To calculate distances between the 7 Siyi dialect points (Taishan, Kaiping, Heshan, Jiangmen, Xinhui, Enping and Doumen; see Fig 1), the number of segmental differences (i.e. segmental Levenshtein distances) between the lexical items among the dialect points were normalized (Heeringa 2004), and fed into a hierarchical cluster model and a non-metric multidimensional scaling (MDS) model to visualize the relationship between the 7 Siyi dialect points.



Results and Discussion: Both the cluster model (Fig 2) and the MDS model (Fig 3) suggest that Heshan is most different from other Siyi dialects. This can be explained by Heshan's proximity to and contacts with Guangzhou, the capital city of Guangdong Province, where the more prestigious Standard Cantonese is spoken. For the remaining 6 Siyi core members, the cluster model divides them into two major groups: Xinhui, Jiangmen and Doumen in one cluster, and Enping, Taishan and Kaiping in another. The MDS configuration suggests that variation within Siyi dialect can be modeled in two dimensions. Dimension 1 represents Heshan's linguistic distance from other Siyi members. Dimension 2, on the other hand, can be interpreted as the geographical location of 7 Siyi dialect points. Taishan, Kaiping and Enping are on one end of Dimension 2 and located in the southwestern Siyi area; while the other 4 dialect points are on the other end of Dimension 2 and located in the northeastern Siyi area. Overall the current study not only quantifies the affinities among 7 Siyi dialect points, but also sheds new lights on the history of Siyi area, that both geographical and social aspects contribute to its internal variation.

Reference:

- Gan, Yu'en. 2002. Guangdong Siyi Fangyan Yufa Yanjiu. Guangzhou, Guangdong: Jinan University dissertation.
- Heeringa, Wilbert. 2004. Measuring dialect pronunciation differences using Levenshtein Distance. Groningen: Rijksuniversiteit Groningen dissertation.
- Tong, Choi-Lan. Guangdong Xinhui Siqianhua Yinyun Yanjiu. Guangzhou, Guangdong: Jinan University dissertation.
- Zhan, Bohui; and Yat-Shing Cheung. 1988. A survey of dialects in the Pearl River Delta vol.2: Comparative lexicon. Guangdong, China: People's Publishing House of Guangdong.