

Aggregate Comparison of Hamgyeong Dialect and *Koryo Mar*

A large proportion of the research pertaining to the varieties of Korean spoken by the diaspora in Central Asia – *Koryo Mar* – has been concerned with establishing its affinity with the varieties of Korean spoken on the Korean Peninsula, most especially those spoken in the area in which the seed community of the contemporary Central Asian diaspora originated in the far North-east of the peninsula.

There is a consensus that *Koryo Mar* bears a high degree of similarity to these varieties, most notably those of the so-called Yukchin dialect zone and select areas of the North Hamgyeong Province dialect zone (c.f. Kho 1987; Kwak 1987; King 1987; King and Yeon 1992; Pak 2005; Kwon 2010 etc.). This observation is made on the basis of a limited number of highly specific linguistic features. While sharing such features as the characteristic Yukchin sentence-final, formal predicate endings (e.g. declarative-*(su)kkwuma*; interrogative - *(su)mtwu*) or the invariant form of the subject particle –*i* strongly suggests close affinity between the transplanted varieties and those of the North East of the Korean peninsula, it does not allow us to draw firm conclusions about their global similarity.

Thus, we suggest an alternative means of examining the relationship between these varieties based on techniques drawn from the linguistic sub-discipline of dialectometry (c.f. Goebel 2010; Wieling and Nerbonne 2015). In this paper we undertake an aggregate comparison of a large amount of dialect data (over 100 items) collected at numerous sites over both Hamgyeong Province (Shinpei 2009 [1940]) and the *Koryo Mar* speaking area (Kwak 2010).

The paper is structured as follows. We first demonstrate the advantages of aggregate methods by highlighting the role of researcher discretion in the selection of features which are used to determine dialectological affinity. We then go on to detail the limitations of this study and discuss the comparability of the data. Having determined its suitability for our purposes, we present visualisations and discussion of our aggregate analysis. The most significant concrete results which arise from this analysis are a series of dendrograms, pseudo-maps and multi-dimensional scaling plots from which the relative aggregate linguistic (dis)similarity between the forms of language used at each of the 39 survey sites (35 in Hamgyeong Province, 4 in Central Asia) may be inferred.

While this paper broadly supports the earlier findings concerning the affinity between *Koryo Mar* and the varieties of the North East of the Korean Peninsula it provides the most nuanced picture of this relationship to date in addition to further promoting empirical, data-driven approaches to the analysis of linguistic variation in Korean.

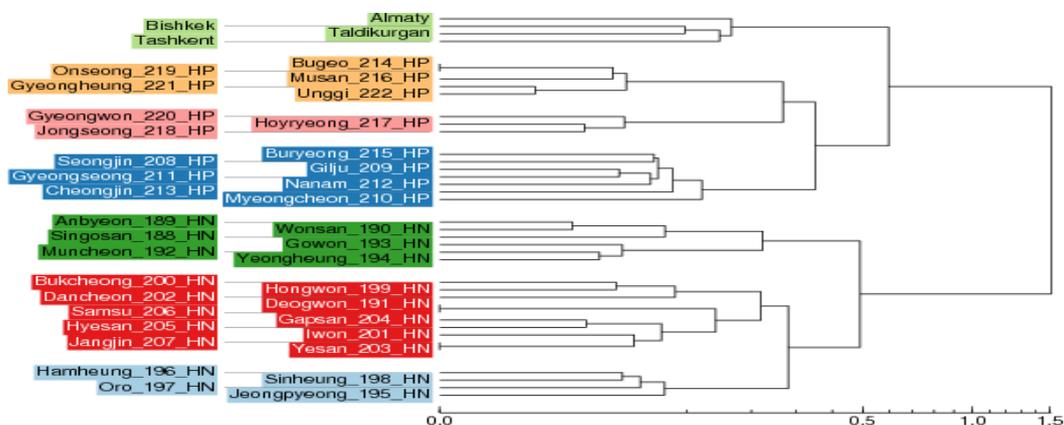


Figure 1: Dendrogram showing Preliminary Clustering of Transplanted and North Eastern Peninsula varieties of Korean.

References

- Goebel, H. (2010). Dialectometry: Theoretical Prerequisites, Practical Problems, and Concrete Applications (mainly with examples drawn from the 'Atlas Linguistique de le France', 1902-1910). *Dialectologia* Special Issue 1, 63-77
- Kho Songmoo (1987). *Koreans in Soviet Central Asia*. Helsinki: Finnish Oriental Society.
- King, R. (1987). An Introduction to Soviet Korean. *Language Research* 23, (2), 233 – 277
- King, R and Yeon Jaehoon (1992). Cwungang asia hanintuluy ene – kolyemal [Koryo Mar – The Language of Central Asia's Koreans]. *Hangul* 217, 83 – 134
- Kwak Chung-gu (1987). Nohan hoyhwawa hampuk kyenghung pangen [Russian Korean Speech and Northern Hamgyeong Kyenghung Dialect], *Cintan Hakpo*, 62, 79-124
- (2010b). *2010nyento kwukoy ciptan icwu hanmincokuy ciyeke cosa 2: kwukoy 5 cicem ehuy, umun, munpep thonhap calyucip, cosa ciyek, cwunkwuk killimseng hwunchwunsi, khacahusuthan almathi, khacahusuthan thaltikhurukan, wucupaykhisuthan thasyukeynthu, khilikisusuthan pisyukheyku* [The 2010 Dialect Survey of Overseas Korean Communities 2: combined collection of resources from five sites including vocabulary, phonology and grammar. Survey sites PRC, Jilin Province, Hoelyongbong; Kazakhstan, Almaty; Kazakhstan, Talkdikurgan; Uzbekistan, Tashkent; Kyrgyzstan, Bishkek]. Seoul: Kwuklip Kwukewen.
- Kwon Jae-il (2010). *Cwungangasia kolyemaluy munpep* [A Grammar of Central Asian Koryo Mar]. Seoul: Seoultaehakkyo chwulphanmunhwawen.
- Pak, N.S. (2005). *Koreyskiy Yazik v Kazakhstane: Problemy I Perspektivy* [The Korean Language in Kazakhstan: Problems and Perspectives] Almaty: Kazakh Ministry for Science and Education.
- Shinpei Ogura (2009). *Cosenepangensacen* [Korean Dialect Dictionary]. Seoul: - Hankwukmunhwasa.
- Wieling, M. and Nerbonne, J. (2015). Advances in Dialectometry. *Annual Review of Linguistics* 1, 243-264