Cross-language dynamics of Okinawa, Japan

Scott A. Hale[1*]
1 Oxford Internet Institute, University of Oxford, UK

Internet communications are highly structured by language and geography (Liao and Petzold, 2013; Hale, 2012; Eleta and Golbeck, 2012; Hong, Convertino, and Chi, 2011; Hecht and Gergle, 2010). While much of this research has identified differences between groups of users on the basis of language or location, little research has investigated the processes by which content is shared between speakers of different languages or examined the impact of such sharing. This research examines such cross-language sharing and its effects in-depth through a mix of quantitative and qualitative methods analyzing the interactions between English and Japanese speakers editing content about Okinawa, Japan, on Wikipedia.

The smaller region permits a more detailed examination of the content shared across languages than previous studies, which have tended to examine global metrics. The research analyzes the history of all Wikipedia articles linking to the "Okinawa" article in English or Japanese. The research identifies the articles appearing in only one of the two language editions (e.g., the crash of a US helicopter into Okinawa International University in 2004 is only covered in Japanese). Article coverage between the two editions is analyzed geographically with geotagged articles and categorically by examining the distribution of articles across categories in the encyclopedia.

As an indicator of the similarity of articles in different editions, the research examines the percentage of images as well as hyperlinks to external sources that are found in common between the two editions. The research addresses what content is shared across the two languages and the characteristics of the users responsible for that sharing.

References:
Eleta, I., & Golbeck, J. (2012). Bridging languages in social networks: How multilingual users of Twitter connect language communities. Proceedings of the American Society for Information Science and Technology, 49(1), 1-4.

Hale, S. A. (2012). Net increase? Cross-lingual linking in the blogosphere. Journal of Computer-Mediated Communication, 17(2), 135-151.

Hecht, B., & Gergle, D. (2010). The Tower of Babel meets Web 2.0: User-generated content and its applications in a multilingual context. Proceedings of the 28th International Conference on Human Factors in Computing Systems (pp. 291-300). New York, NY, USA: ACM.

Hong, L., Convertino, G., & Chi, E. (2011). Language matters in Twitter: A large scale study. International AAAI Conference on Weblogs and Social Media (pp. 518-521).

Liao, H.-T. (2013, August 5–7). How do Baidu Baike and Chinese Wikipedia filter contribution? A case study of network gatekeeping. Wikisym 2013: The Joint International Symposium on Open Collaboration, Hong Kong.