Jinyun Ke(1), Feng Wang(1), and Christophe Coupé(2)

(1) Language Engineering Laboratory, Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China (2) Laboratoire Dynamique du Langage, Université Lyon 2 Louis Lumière, Lyon, France.

jyke@ee.cityu.edu.hk

"The Rise and Fall of Homophones: a Window to Language Evolution"

Homophony is believed to be a ubiquitous phenomenon in all human languages. However, its existence is considered as non-adaptive in language systems, as it brings ambiguity in communication, though unnoticeable sometime (Wang, 2001). These observations lead us to investigate the composition and evolution of homophones. In this paper, we report a comprehensive analysis of the distributional characteristics of homophones from a cross-linguistic point of view. We also present studies on the evolution of homophones from empirical analysis as well as computational simulations.

We analyzed homophony by comparing six Chinese dialects, as well as English and German. Languages with larger phonological inventories and more complex syllable structures are found to have fewer homophones. In particular, for the Chinese dialects, a very clear correlation appears between the phonological complexity and the degree of disyllabification. Disyllabification may be a result of homophony avoidance. Furthermore, languages which have more morphological constraints have larger tolerance for homophones. Regarding distribution, we find that homophones within the same grammatical class are much less numerous than those dispersed in different classes. For homophones in the same grammatical class, further distinctions can be made between the words according to their usage frequency.

Homophones in a language evolve continuously. While new homophones emerge from various sources, including sound change, imperfect learning and borrowing, old homophones may die out as their usage decreases. The disappearance of words homophonous to taboos is a typical example of homophone demise (Bloomfield, 1933). However, the process is so slow and elusive that it can be observed only at a long time scale.

In this study, we developed a multi-agent system to model the rise and fall of homophones. Agents' vocabulary is represented in a novel way: two associated self-organizing maps represent phonological and semantic forms respectively. The communication among agents is modeled by iterative one-to-one speaking and listening interactions. It is demonstrated that homophones emerge due to constraints in production and perception which are introduced in the phonetic space (Nowak et al., 1999), and homophones disappear due to communication failures. These processes lead to continuous internal re-organization in the agents and consequent sound changes. We observe an interesting periodic pattern of chained emergence and disappearance of homophones. The current model will benefit from integrating insights from language acquisition and overlapping generations in a more realistic way, which may reveal further intrinsic dynamics of language evolution.

We conclude that the existing organized structures in homophones are emergent characteristics due to the self-organization in a system. Homophones undergo incessant evolution, providing an informative window for zooming in on the picture of language evolution.

References:

1. Bloomfield, Leonard. 1933(1984). *Language*. The University Chicago Press.

2. Wang, William S-Y. 2001. Language evolution and leaky grammars. Paper presented at the Workshop on Language Learning and Evolution. Institute for Advanced Study, Princeton.

3. Nowak, Martin, Joshua B. Plotkin and David C. Krakauer. 1999. The evolutionary language game. *Journal of Theoretical Biology*, 200:147-162.