Grapheme-to-Phoneme Conversion for Korean Point-of-Interest Data Using Prosodic Words

Sunhee Kim (sunhkim@snu.ac.kr)
Center for Humanities and Information, Seoul National University
San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, Korea

Je Hun Jeon (jhjeon@snu.ac.kr)
Center for Humanities and Information, Seoul National University
San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, Korea

Minsoo Na (dix39@snu.ac.kr)
Interdisciplinary Program in Cognitive Science, Seoul National University
San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, Korea

Minhwa Chung (mchung@snu.ac.kr)
Department of Linguistics, Seoul National University
San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, Korea

Keywords: Grapheme-to-Phoneme conversion; Prosodic word; Compound; Proper name; Point-of-Interest data; Phonology

Point-of-Interest (POI) data are location enriched word data, which consist of personal names, place names including administrative districts, and names of business and recreational locations. They are essential to the development of applications such as Telematics, wireless location-based services (LBS), and tourism and other business applications. Being a major component of Text-to-Speech (TTS) systems and automatic speech recognition (ASR) systems, a Grapheme-to-Phoneme (G2P) conversion system maps the orthography to its corresponding sequence of phonemes. When POI data are converted into corresponding pronunciations, two characteristics of the data need to be considered. First, they are proper names, which are known to be closely related to irregular pronunciations in Korean. Second, most POI data are compound or complex words consisting of more than two words. A constituent of a POI word belongs to many other words. Therefore, it is a prerequisite to separate each word into its constituents in order to derive its pronunciation which includes irregular pronunciations.

In this paper, we propose a G2P conversion system that detects irregular pronunciations for Korean POI data, by adopting the notion of prosodic word based on Prosodic Phonology (Selkirk 1984). The performance of the proposed system depends on the effective detection of irregular pronunciation words. According to our proposed system, all the words in the POI data were broken down into prosodic words, which were trained by category (regular, irregular or referential). The trained words were then used to create a dictionary for each category. By adopting this proposed system, irregular pronunciation words could be detected with increased efficiency using a far lesser number of words than other methods. For detecting words of irregular pronunciation, only 22.2% of input data on average were required for manual review for our proposed system, compared to 87% required when using the exceptions detection approach without using prosodic word as unit, and to 100% required when using the regular data-driven approach. The results of our experiment showed that 11.43% of 300k POI data were detected as being words of irregular pronunciation. Compared to the 6.67% obtained in a text corpus, POI data consisting of proper names comprise a much higher percentage of irregular pronunciations than ordinary text corpus.

By combining phonology with computational modeling, the proposed system not only provides enhanced performance but also, assuming that phonology is a subsystem of the mind/brain (Durand & Laks 2002) or a branch of the study of mental representation, the psychology of mind (Hale & Reiss 2000), reflects the same method by which a speaker derives pronunciation of morphologically complex Korean proper names..

References

Durand, J. & B. Laks (2002). *Phonetics, Phonology, and Cognition*, Oxford University Press

Hale, M. & C. Reiss (2000). Phonology as Cognition, *Phonological Knowledge: Conceptual and Empirical Foundations*, N. Burton-Roberts, Philip Carr & Gerard Docherty, eds. Oxford University Press. pp.161-184.

Selkirk, E. (1984). Phonology and Syntax: The Relation Between Sound and Structure, Cambridge: MIT Press.