Available online at www.sciencedirect.com

ScienceDirect

Lingua 159 (2015) 70–92

Lingua

www.elsevier.com/locate/lingua

When in doubt, read the instructions: Orthographic effects in loanword adaptation

Robert Daland [a], Mira Oh [b,*], Syejeong Kim [b]

[a] Department of Linguistics, UCLA, United States
[b] Department of English, Chonnam National University, Republic of Korea

Received 1 February 2014; received in revised form 25 February 2015; accepted 3 March 2015
Available online 15 April 2015

Abstract

Loanword adaptation has yielded many insights into the relationship between speech perception and the phonological grammar. Evidence is now mounting that orthographic effects on loanword adaptation may be more prevalent than was once thought (cf. Paradis and LaCharité, 2002; Vendelin and Peperkamp, 2006), partially obscuring phonological effects. This paper investigates orthographic effects in the adaptation of vowels of English words loaned into Korean. Experiment I uses information-theoretic statistics, called the *orthographic* and *perceptual information gains*, to estimate a lower bound on the contribution of orthography and perception to vowel adaptation. The results suggest that orthography contributes more to the adaptation of unstressed vowels, while perception contributes more to the adaptation of stressed vowels. Experiment II considers the adaptation of the /ɛ/∼/æ/ contrast; these vowels have merged recently in Korean although the orthographic distinction is maintained. The paper concludes by proposing the Perceptual Uncertainty Hypothesis: source-loan orthographic alignment plays the greatest role in constraining loanword adaptation when phonological parsing in the borrowing language is underdetermined by perceptual factors alone.

Keywords: Loanword phonology; Speech perception; Korean; Information theory

1. Introduction

Loanword adaptation provides a unique window onto the relation between speech perception and the phonological grammar. For example, loanword adaptation data have facilitated the discovery of phonetic factors that contribute to phoneme identification in native and non-native speech perception (Davidson, 2007; Kang, 2003a,b). Similarly, the desire to formally account for loanword adaptation has supported key tenets of constraint-based phonology, and spurred formal innovations pertinent to L1 phonology (Boersma and Hamann, 2009; Shinohara, 2004).

Researchers agree that the adaptation process yields forms which are generally similar to the source form, while also generally conforming to phonotactic properties of the borrowing language (Iverson and Lee, 2006; Kang, 2010; LaCharité and Paradis, 2002; Peperkamp et al., 2008; Shinohara, 2004; Silverman, 1992; Yip, 1993). Debate has centered on the extent to which loanword adaptation reflects direct perception versus phonologically informed adaptation. It now seems clear that certain aspects of loanword adaptation reflect perceptual distortions that derive from the structure of borrowers' native language (e.g. Dupoux et al., 1999; Kabak and Idsardi, 2007; Boersma and Hamann, 2009), while other aspects of

* Corresponding author.
E-mail address: mroh@chonnam.ac.kr (M. Oh).

http://dx.doi.org/10.1016/j.lingua.2015.03.002
0024-3841/

loanword adaptation reflect post-perceptual repair of a form that violates L1 phonotactics (Shinohara, 2004; Kabak and Idsardi, 2007; for an analogous point in L2 phonology see Broselow et al., 1998). Adjudicating this tension has paid off theoretically by increasing our understanding of how speech perception shapes synchronic phonology, and vice versa.

Much of this progress involved the use of loanword corpora (Kang, 2010; Paradis and LaCharité, 2002; Peperkamp et al., 2008; Shinohara, 2004; Silverman, 1992; Yip, 1993). Corpus studies are methodologically attractive because loanword corpora are comparatively easy to collect, analyze, and share between research teams; thus, the time required to do a study is fairly small, the number of items is large, and replicability is good. On the other hand, the processes by which loanwords are introduced and standardized are not fully understood, and are likely subject to a wide variety of extra-perceptual influences. An especially salient source of 'contamination' is orthography -- adapters may be guided by aspects of the source form's orthographic representation rather than by its phonological (and perceptual) properties. LaCharité and Paradis (2005) conducted an influential study of French loanword phonology, in which they concluded that English orthography affected the French adaptation in a negligible proportion of forms. This view appears to reflect the conventional wisdom, as the majority of papers on loanword adaptation that we are familiar with do not make more than passing reference to orthography. One notable exception is Vendelin and Peperkamp (2006), who make orthography the focus of their study. They conducted an ''online adaptation'' study with French listeners, and found that the inclusion of English orthography conditioned how French listeners adapted the forms.

This paper follows in the footsteps of Vendelin and Peperkamp by studying the role of orthography in loanword adaptation, but with a completely different language: Korean. The research question this paper takes up is driven by the authors' informal observation that the adaptation of English unstressed vowels into Korean appears to be guided by orthography. Experiment I is a loanword corpus experiment, designed to assess the relative effects of orthography versus perceptual factors in vowel adaptation. The methodology uses concepts and statistics from information theory to partially separate orthographic and perceptual factors. Experiment Ia demonstrates that both orthography and perception play a role in vowel adaptation; Experiment Ib provides suggestive evidence that the role of orthography is greater for unstressed vowels, while perception may play a stronger role in the adaptation of stressed vowels. Experiment II considers the adaptation of the English /ɛ/-/æ/ contrast (typically adapted into Korean orthography as ㅇㅔ and ㅇㅐ respectively). The vowels spelled ㅇㅔ and ㅇㅐ underwent a phonetic merger sometime around the 1950s in the standard dialect of Korean, although the spelling distinction is retained in the orthography (Hong, 1988; Choi, 2002; Kim, 2000; Chung, 2002). Experiment II, an online adaptation study with nonce words, shows that Koreans do not exhibit distinct adaptation patterns for English [ɛ]-[æ] when exposed to the auditory forms alone (suggesting they are unable to distinguish these vowels), but when orthography is included, the adaptation patterns shift toward what was found in the corpus study. The paper concludes by introducing the *Perceptual Uncertainty Hypothesis*, that source-loan orthographic alignment constrains the phonological and orthographic parse assigned to a loan, and that this constraint is supplementary to perceptual adaptation, so that the strongest orthographic effects will be observed when perceptual factors alone underdetermine adaptation.

1.1. Conventions

We employ several conventions in this paper. In cases where the language of a form might be ambiguous, we supply a subscript E after English forms, and a subscript K after Korean forms. We use tilde-angle bracket sequence to indicate loanword adaptation, and a hyphen-angle bracket sequence to indicate native phonological mappings, e.g. $word_E \sim\!> /w\Lambda ti/_K$ --> $[w\Lambda di]_K$. Tense obstruents are denoted with a '*' diacritic. Finally, while the Korean vowels ㅇㅔ and ㅇㅐ are often transcribed as /e/ and /ɛ/ respectively, we symbolize both as /ɛ/ to reflect the vowel merger (e.g. Hwang and Moon, 2005), except in Experiment II, where we nonstandardly transcribe ㅇㅔ as /ɛ/ and ㅇㅐ as /æ/. This transcription reflects the preferred adaptation for the corresponding English vowels; we do this so as to minimize the cognitive load on readers not already familiar with Korean phonology.

1.2. Korean phonology and orthography

Korean (Ahn, 1998; Sohn, 1999; Shin et al., 2012) includes 7 monophthongal vowels /i, ɨ, u, ɛ, ʌ, o, a/, and various diphthongs. The sonorant consonant inventory includes front and back glides /j, w/; a single liquid /l/ (realized as a tap [ɻ] in onset position), and nasals at the major places of articulation /m, n, ŋ/. Stops contrast for four places of articulation (labial, denti-alveolar, alveo-palatal, and velar); the alveo-palatal stops are affricated. The stops exhibit a typologically unusual three-way laryngeal contrast: plain/lax /p, t, c, k/, aspirated /p^h, t^h, c^h, k^h/, and tense/fortis /p^*, t^*, c^*, k^*/. There are lax and tense coronal fricatives /s, s^*/ (which alveopalatalize before [j, i], yielding [ɕ, $ɕ^*$]), and a glottal fricative /h/ that is subject to various lenition and coalescence processes.

The basic syllable structure is (C)(G)V(C). In the coda position, only the liquid, nasals, and plain oral stops may occur, a restriction that is enforced by active laryngeal neutralization (tense and aspirated stops become plain) and active place/manner neutralization (alveo-palatal stops and all fricatives become plain denti-alveolar stops). Underlying laryngeal and