

Korean vowel merger effects on English loanword adaptation

Stanley Nam (University of British Columbia, Canada)



Introduction

Prescriptive norm

English→	Korean	Example
[æ]	ㅏ ay	damage → 대미지
[ɛ]	ㅓ ey	penalty → 페널티



Speakers' vowel system

- ㅏ ay–ㅓ ey merger [1–3]
- No reliable perception of the source vowel contrast [3, 4]

- The mismatch raises the question of whether loanword adaptations **preserve** the **prescribed contrast** in contemporary usage.
- This study examines whether and how English loanword adaptations **depart from** the prescribed ㅏ ay–ㅓ ey correspondences.
- The analysis draws on **novel data** from a corpus of Korean-language posts from **Bluesky**.

Research question and hypothesis

Research Question

- Does the ㅏ ay–ㅓ ey merger affect English loanword adaptation, and if so, how?

Hypothesis

- Loanword adaptations will not consistently preserve the prescribed contrast.

Methods

Overview

- Bluesky is a social media platform like Twitter.
- A corpus of ~11M Korean-language Bluesky posts was collected from 7,183 accounts.
- Occurrences of possible adaptation forms were surveyed in 1,070 accounts.
- Confusion rates (CR) were calculated for individual loanwords.
- Additionally, individual target ㅏ /ㅓ tokens were analyzed for whether confusion occurred and, if so, in which direction.

Corpus (Korean-language Bluesky posts)

- Step 1:** Identified 7,183 accounts that posted in Korean during a 24-hour window.
- Step 2:** Collected up to 10,000 posts from each account retrospectively.
- Corpus size and analyzed subset**
For this preliminary study, only accounts with 1K–3K posts were included in the analyses reported here (17.64% of all collected posts).

	Collected	Analyzed
Accounts	7,183	1,070
Posts	10,864,876	1,916,926

Target loanwords (N=4,591)

- Raw data:**
12,212 loanword nouns in *Standard Korean Language Dictionary* as compiled by [5].
- Custom Python script:**
Extracted unique ㅏ /ㅓ -containing entries from the raw data, while excluding ㅓㅓ sequences representing [eɪ].

Methods (cont'd)

Possible adaptations and confusion rate

- ㅏ /ㅓ spelling combinations were generated for each word and labelled Form₁, ... Form_n.
- The corpus freq. of Form_i was denoted by f_i .
- Let p index the prescribed form.
The confusion rate was calculated as:

$$\text{Confusion rate} = 1 - \frac{f_p}{\sum_{i=1}^n f_i}$$

e.g., 'damage' [dæmidʒ], ㅏ ay is prescribed.
→ Form₁: 대미지 (with ㅏ ay), $f_1 = 190$
→ Form₂: 데미지 (with ㅓ ey), $f_2 = 356$

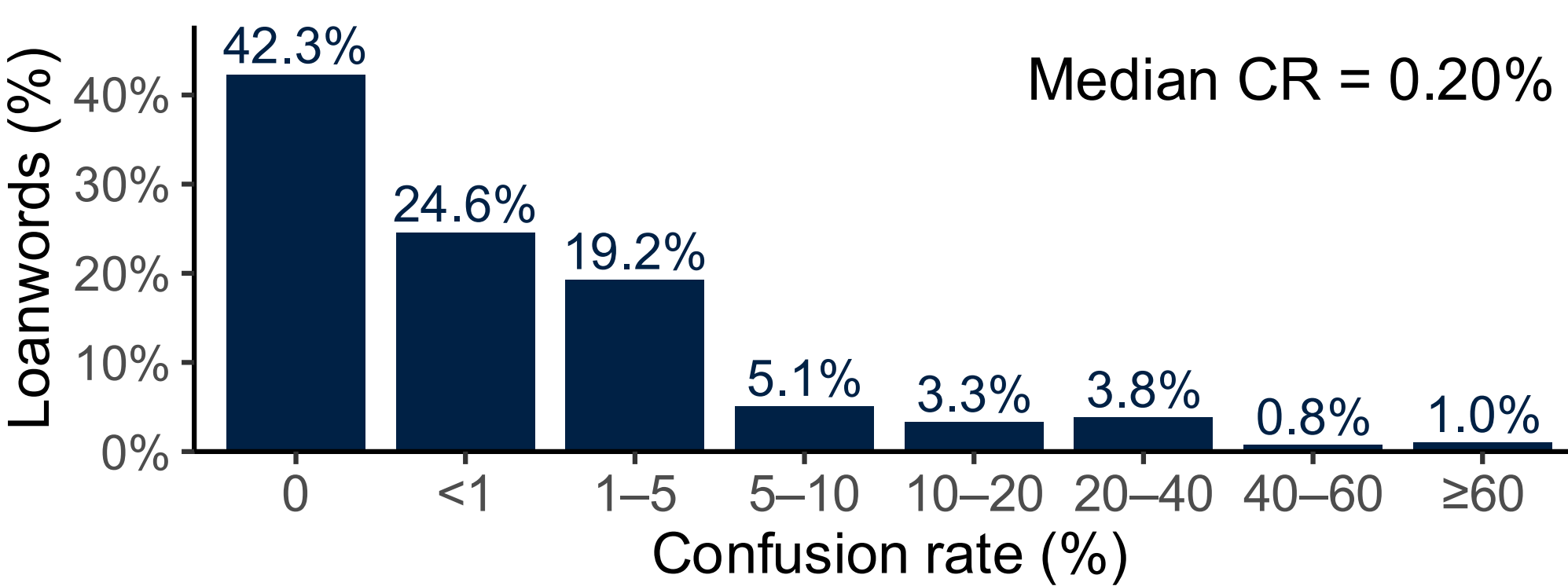
$$\text{Confusion rate} = 1 - \frac{190}{190+356} = 0.6520$$

Data filtering (N=395)

- Only words meeting both thresholds were analyzed: ≥ 10 in [6] and ≥ 30 in the corpus.

Results

Word-level confusion patterns (N=395)



Loanwords with high confusion rates

Entry	Prescribed	Variant	CR (%)
leisure	레저 ㅓ ey	래저 ㅏ ay	84.12
damage	데미지 ㅏ ay	데미지 ㅓ ey	65.20
penalty	페널티 ㅓ ey	패널티 ㅏ ay	63.37
rally	랠리 ㅏ ay	렐리 ㅓ ey	60.78

Results (cont'd)

Confusion matrix (target-site token counts)

Prescr.	Obs. ㅏ ay	Obs. ㅓ ey	Total
ㅏ ay	89,603	4,669	94,272
ㅓ ey	2,862	148,369	151,231
Total	92,465	153,038	245,503

- ㅏ ay → ㅓ ey: $4,669 \div 94,272 = 4.95\%$
- ㅓ ey → ㅏ ay: $2,862 \div 151,231 = 1.89\%$
- Confusion direction (N=7,531 confusions):
ㅏ ay → ㅓ ey: 62%; ㅓ ey → ㅏ ay: 38%

Discussion and conclusion

- This preliminary study finds **little evidence** that the ㅏ ay–ㅓ ey merger affects adaptation. ↔ studies showing target-language changes influence adaptation [6–8].
- Confusion was concentrated** in a small subset of English loanwords.
→ The variation appears item-specific, rather than systematic.
- Confusion was asymmetric:**
ㅏ ay → ㅓ ey had a 3.06-pp higher CR and accounted for 62.00% of all vowel confusions.
- The asymmetry may reflect a default bias based on prescribed-vowel frequency: ㅓ ey was 1.60 times as frequent as ㅏ ay.
- Future work** can examine lexical frequency, user dispersion, and source spelling [9].



References

- [1] Hong, Y. 1987. The e/a merger in modern Seoul Korean. In S. Kuno, I.-H. Lee, J. Whitman, S.-Y. Bak, & Y.-S. Kang (Eds.), *Harvard studies in Korean linguistics II* (pp. 367–376). Hanjin Publishing Company.
- [2] Lee, Y.-C., & S. Cho. 2021. The merger of /e/ and /ɛ/ in Seoul Korean tracing speakers born from the 1930s to the 1960s. *Journal of Linguistic Studies* 61: 1–16.
- [3] Eychemme, J., & T.-Y. Jang. 2015. On the merger of Korean mid front vowels: Phonetic and phonological evidence. *Journal of the Korean Society of Speech Sciences* 7(2): 119–129.
- [4] Ingram, J. C. L., & S.-G. Park. 1997. Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics* 25(3): 343–370.
- [5] Park, N. 2020. *Statistical learning of Korean phonotactic constraints and well-formedness judgments*. PhD dissertation, Seoul National University.
- [6] Kang, Y. 2008. Tensification of voiced stops in English loanwords in Korean. In S. Kuno, I.-H. Lee, J. Whitman, J. Maling, Y.-S. Kang, P. Sells, H.-S. Sohn, & Y. Jang (Eds.), *Harvard studies in Korean linguistics XII* (pp. 179–192). Hanjin Publishing Company.
- [7] Nam, S. 2021. The adaptation of English word-initial voiced stops in Korean: A diachronic approach. *Studies in Phonetics, Phonology, and Morphology* 27(1): 3–25.
- [8] Lee, H. 2020. An acoustic comparison of phonetic realization between South Kyungsang Korean-speaking children and adults' English loanword production. *Studies in Phonetics, Phonology, and Morphology* 26(2): 303–320.
- [9] Vendelin, L., & S. Peperkamp. 2006. The influence of orthography on loanword adaptations. *Lingua* 116(7): 996–1007.