eScholarship
Japanese/Korean Linguistics

Title

Is Korean Sign Language (KSL) Place Name Initial Morpheme Skipping Triggered by Semantics or Phonology?

Permalink

https://escholarship.org/uc/item/2546q3qd

Journal

Japanese/Korean Linguistics, 32(1)

Author

George, Johnny

Publication Date

2026-07-31

DOI

10.5070/J7.66606

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at https://creativecommons.org/licenses/by-nc-nd/4.0/

eScholarship.org Powered by the California Digital Library
University of California

Is Korean Sign Language (KSL) Place Name Initial Morpheme Skipping Triggered by Semantics or Phonology?*

JOHNNY GEORGE
Meiji University

1 Introduction

An investigation of 106 Gwangju Korean Sign Language (KSL) toponyms by Ji (2022) describes most names as bimorphemic compounds indexed to the source Korean names, thus loan translations. Roughly a quarter of the KSL names simply represent the initial source morpheme semantics; however, four names represent only the second source morpheme semantics, apparently flouting a constraint for monomorphemic outputs that requires indexing the initial source morpheme semantics. At JK31, two hypotheses were proposed to explain preference for the second morpheme in the KSL toponym: 1) the greater semantic salience of the second source morpheme makes it favored over the initial source morpheme; and 2) an easily assimilated KSL Hangul character sign initialization indexing the initial source morpheme satisfies a general constraint supporting indexing any feature of the initial source morpheme, thus allowing the second morpheme to appear alone while preventing ambiguity. This work examines these two hypotheses.

* Special thanks to Youngju Choi and the Chosun Language Convergence Research Institute researchers for providing KSL toponym references and helpful comments. The author bears sole responsibility for errors or omissions in this work.

Japanese/Korean Linguistics 32.
Edited by Youngdong Cho, Young-Hoon Kim, and John Whitman.
Copyright © 2026, Johnny George.

2 Sign Language Toponyms and Semantic Transparency

Various studies on sign language toponyms posit that sign language toponym outputs tend to index spoken source toponym morphemes that have transparent semantics (George 2025a; McKee and Mckee 2000; Sutton-Spence and Woll 1999; Peng and Clouse 1977). While most studies determine semantic transparency through introspection, ideally empirical tests would measure semantic transparency. For instance, in an examination of JSL toponyms, George (2025a) defines relative semantic transparency based on the grade level at which children learn a particular Japanese character and finds that earlier learned characters appearing in Japanese source toponyms tend to have their semantics represented in JSL toponyms outputs at a greater proportion than those learned later in life.

This study weighs the semantic salience of source toponyms through proxies for two characteristics of prototypical words based on criteria discussed in Sutrop (2000), psychological relevance, and grounding in a fundamental concept. Sutrop (2000) gives a cognitive grounded basis for identifying basic words in a lexicon. Basic words tend to serve as prototypes, which easily and initially come to mind, and have uses across a range of contexts. Widely used words have psychological relevance for language users as they form the basis of communication in daily life. If KSL toponyms tend to index semantically salient morphemes from source Korean toponyms, then morpheme skipping might receive an account from the identification of basic Korean morphemes.

3 Method

To test the influence of an input semantic constraint versus a phonological output constraint, this study examined the 106 KSL metro names from *Gwangju Metropolitan City Sign Language Place Names* (Gwangju Metropolitan Association of the Deaf, 2021). A spreadsheet of all source toponym *Hanja* (Chinese Characters) and *Hangul* tracked the KSL toponym outputs to analyze the distribution of *Hanja* and *Hangul* indices[1]. Following the procedure from George (2022) for JSL, each individual *Hanja* and *Hangul* represented in the source toponyms had its source input to KSL output ratio probability measured. This ratio appears labeled as Character Preservation Frequency in the tables below.

Along with measuring the Hanja distribution, this study weighs semantic salience by examining proxies for two characteristics of prototypical words based on criteria discussed in Sutrop (2000), psychological relevance and grounding in a fundamental concept. *KoTenTen18* Corpus (Sketch Engine 2028) word counts of the source morphemes' Hanja characters represent psychological relevance, grounded in the hypothesis that high word frequency increases familiarity and accessibility. Appearance in the Holman et al. (2008) ranked Swadesh list represents grounding in a fundamental concept, with the proposition that such basic words appropriately represent basic facets of human experience.

[1] KSL has signs that trace Sinographs, so for character-to-KSL sign comparisons *Hanja* character representations appear as referenced in Gwangju Metropolitan Association of the Deaf (2021). This paper does not claim *Hanja* as underlying mental representations. See work such as Kim (2019) for a discussion of the degree of *Hanja* awareness among Korean L1 speakers.

4 The Character Preservation Frequency of Hanja and Hangul

With respect to the distribution of KSL *Gwangju City* toponym types, around a quarter of the names are monomorphemic while the remainder emerge as polymorphemic (Table 1). An overwhelming majority (93%) are exonyms, names based on the Korean source toponym, while 5% are endonyms, original sign names independent of the source Korean name, and 2% are endonymic-exonymic outputs. Seventeen toponyms have Hangul initialization[2] with roughly half initializing monomorphemic outputs and half in polymorphemic outputs.

	Mono	**Bi**	**Tri**	**Total**
All Names	27	70	3	100
Exonyms	24	66	3	93
Endonyms	3	2	0	5
EndoExo	0	2	0	2
Initialized	8	8	1	17

Table 1: Distribution of KSL *Gwangju City* toponym types

The Character Preservation Frequency for the ten *Hanja* source characters with the greatest output frequencies appear below (Table 2). For instance, ten toponyms contain the character sign 月 월 'moon' and KSL toponyms index them in ten out of ten source names. Although lacking statistical significance, the distribution offers important clues of where to find patterns in any future investigation of a larger database of KSL toponyms. Several input characters appear 100% of the time in the KSL toponym outputs. Despite the lack of statistical significance, this result suggests that inclusion of a larger collection of KSL toponyms could still result in a salient pattern of representation for certain characters, as found for JSL toponyms (George 2022).

[2] Initialized toponyms are referred to by George (2025b) in translation from Ji (2022) as "blends". This work will refer to these as initialized toponyms to remain consistent with prior work on sign toponyms.

	Char		Type	n=	k=	p=	P(k)= [a]	sd=	mean=
1	**月**	**월**	char	**10**	**10**	0.80105	**0.1088**	1.2624	8.0105
2	**山**[3]	**산**	char	**8**	**8**	0.80105	**0.1695**	1.1291	6.4084
3	林	림		7	6	0.80105	0.368	1.0562	5.6073
4	岩	암	char	5	5	0.80105	0.3298	0.8927	4.0052
5	東	동	cardinal	4	4	0.80105	0.4117	0.7984	3.2042
6	雲	운		4	4	0.80105	0.4117	0.7984	3.2042
7	谷	곡		5	4	0.80105	0.4096	0.8927	4.0052
8	**龍**	**용**	char	**4**	**4**	0.80105	0.4117	0.7984	3.2042
9	大	대		4	3	0.80105	0.4091	0.7984	3.2042
10	孝	효		3	3	0.80105	0.514	0.6915	2.4031

Table 2[4]: *Hanja* Character preservation frequency

To examine the pattern of initialization across the KSL inventory, a spreadsheet determined the character preservation frequency for each source name *Hangul* initial phoneme represented in all KSL name outputs (Table 3). Two particular KSL *Hangul* characters emerged as onset indexes to greater degree relative to other input Hangul segments, ㄷ t (4/10) p=.067 and ㅂ p (3/7) p=.092. The ease of assimilation of these segments may account for their relatively high distribution.

Statistically notable for a lack of representation, the Korean word initial character ㅇ serves as a null segment that signals a vowel initial word in the orthography; KSL represents it isomorphically, and it only emerges in the dataset to represent an initial [o] phoneme, which serves as the KSL default vowel representation. The Korean toponym dataset consists of 22 vowel initial names, and no other word initial vowel signs are represented in the KSL sign output. Apart from [o]-initialized KSL names, consonants exclusively appear as initializations.

[3] 山 does not include two homophonic variant names, 牛山 우산 *Usan* as 우산 'umbrella' and 山水 산수 *Sansu* as 산수 math, considering those, n=10, k=8 and p<.3015

[4] Notes. Char sign = character sign that maps with some portion of the Hanja orthography. Cardinal = cardinal sign. p=170 out of 193 (80%) character-inputs represented in the KSL toponym outputs. This count does not include any characters that are not semantically represented; therefore, outputs solely indexing phonemes are not included. George (2022) considers p<.02 results as statistically significant; therefore, this study will maintain the same benchmark, as a result, there are no statistically significant patterns of distribution. [a] $P(k) = {}_nC_k\, p^k q^{n-k}$.

			n=	k=	p=	P(k)=	sd=	mean=
1	ㅅ	s	22	2	0.18	0.141395	1.80199889	3.96
2	**ㅇ**		**22**	**1**	0.18	**0.06134598**	1.80199889	3.96
3	**ㅌ**	**t**	**10**	**4**	0.18	**0.06701815**	1.2149074	1.8
4	ㄱ	k	7	1	0.18	0.38304841	1.01646446	1.26
5	ㅈ	c	7	0	0.18	0.24928547	1.01646446	1.26
6	ㅁ	m	7	2	0.18	0.25225139	1.01646446	1.26
7	**ㅂ**	**p**	**7**	**3**	**0.18**	**0.09228709**	1.01646446	1.26
8	ㅎ	h	5	1	0.18	0.40690958	0.85906926	0.9
9	ㅊ	ch	4	1	0.18	0.39698496	0.76837491	0.72
10	ㄴ	n	4	2	0.18	0.13071456	0.76837491	0.72
11	ㅍ	ph	3	1	0.18	0.363096	0.66543219	0.54
12	ㅆ	ss	2	0	0.18	0.6724	0.54332311	0.36

Table 3[5]: *Hangul* Character initialization preservation frequency

5 Initialized Names

In the distribution of the 17 initialized toponyms from the dataset, seven index the semantics of both morphemes of the source toponym (Table 5); four index the semantics of the first morpheme only; and four index the semantics of the second morpheme only (Table 6). Typically, when the semantics of the source toponym is represented either the first or both morphemes emerge in the KSL toponym, so the instances that fail to represent the initial morpheme semantics appear to flout some type of constraint.

1st morpheme	2nd morpheme	Tokens
(initialized)-semantic	(initialized)-semantic	7 (Table 5)
initialized-semantic	——	4
initialized	**semantic**	**4 (Table 6)**
(Initialized)-semantic	initialized	Samto 삼도 (三道)
initialized-semantic	initialized	Hwaceng 화정 (花亭)

Table 4[6]: - Initialization + Semantics of source morphemes distribution

	Hanja	Hangul	Ji Ref #	Romanized	Out1	Out2

[5] $P(k) = {}_nC_k\, p^k q^{n-k}$

[6] Based on counts from the database in this study; it contains slightly different values than Ji (2022).

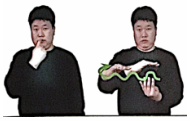
	東明	동명	5	*Tongmyeng*	East Bright	ㄷ-東	明
	治平	치평	27	*Chiphyeng*	Govern Peace	ㅊ-治-平	
	白雲	백운	33	*Paekwun*	White Cloud	白	ㅂ-雲
	大村	대촌	48	*Taychon*	Large Village	ㄷ-大	村
	梅谷	매곡	52	*Maykok*	Plum-Blossom Valley	ㅁ+梅	谷
	道山	도산	75	*Tosan*	Road Mountain	ㄷ-**道**	**山**
	道泉	도천	76	*Tochen*	Road Spring	道	ㄷ-道 +ㄷ-泉

Table 5[7]: Initialized-Semantic + Semantic toponyms

ㅁ-月(m-month)		ㄴ-門(n-door)		ㅅ-龍(s-dragon)		ㅇ-山 (o-mountain)	
16/295	*335/7738*	*97/1272*	*414/6871*	*208/5307*	*369/6257*	*10/125*	*5/67(1233/21K)*
每	月	樓	門	臣	龍	梧	峙
매	월	누	문	신	용	오	치
***m**ay*	*wel*	***l**u*	*mun*	***s**in*	*yong*	***o***	*chi*
‘every’	‘**moon**’	‘building’	‘**door**’	‘vassal’	‘**dragon**’	‘Parasol tree’	‘**peak**’

Table 6: Initialized — 1st consonant initialized + 2nd morpheme semantics indexed (Ji 2022: 175-176)

[7] All sign language images from Gwangju Metropolitan Association of the Deaf (2021).

6 Identifying Semantic Salience

Three factors serve as proxies for prototypical and semantically transparent words with crosslinguistic conceptual grounding. Words are checked for their: character preservation frequency; frequency counts from the *KoTenTen18 Corpus* of Korean words; and appearance in the Swadesh list.

Three of the second position morphemes, 月 월, 山 산, and 龍 용, appear among the class of most represented morphemes. The *Hanja* character preservation frequency chart shows that particular source morphemes emerge 100% of the time, the most represented being: 月 월 *wol* 'moon' (10/10), 山 산 *san* 'mountain' (8/8), 岩 암 *gam* 'rock' (5/5), 東 동 *dong* 'east' (4/4), 雲 운 *un* 'cloud' (4/4) and 龍 용 *yong* 'dragon' (4/4), signaling that these inputs may have semantic salience for KSL toponym creation (Table 2). In contrast, none of the initial morphemes are preserved at a greater rate than the second morpheme, so this pattern suggests that some quality might drive the preference for the indexed morphemes.

Frequency counts for each morpheme from the *KoTenTen18* corpus appear (Table 6). The first number shows the type-frequency of the character as part of a unique word, and the second number provides the token count of all appearances of the word. For instance, "every" appears in 16 different words and a total of 295 times across all appearances within all words. For the four names in the chart, three of the four 2nd position morphemes appear at a greater frequency than the name initial morpheme. The exception, "peak" does not have an exact semantic equivalent in KSL so emerges as the sign "mountain". The frequency of "mountain" in the corpus is 1233 types with 21,000 tokens, so the actual sign output represents the more common morpheme. The KSL outputs appear to favor the more frequently used morphemes. The frequency of the four morphemes suggest that they may have higher psychological relevance than the skipped initial morphemes.

The semantics of each morpheme was compared based on its appearance in the ranked Swadesh list. 月 월 *wol* 'moon' and 山 산 *san* 'mountain', appear in the ranked Swadesh list, suggesting that these represent fundamental concepts.

The findings suggest that semantic salience may partially account for the indexation of the non-initial morpheme semantics. *KoTenTen18 Corpus* type/token word counts show all four indexed morphemes with greater frequency than the skipped initial morpheme, while two of the morphemes, 月 월 *wol* 'moon' and 山 산 *san* 'mountain', appear in the ranked Swadesh list. Based on the criteria, 月 and 山 appear as highly prototypical morphemes, and perhaps 龍 along with 門 to a lesser degree. In contrast, the initial morphemes in this set of words meet none of the semantic salience criteria (Table 7).

Further supporting the significance of semantic salience is the likely lack of semantic transparency for 樓 누 nu 'building', 臣 신 sin 'vassal', and 梧 오 o 'Parasol tree', three words that lack KSL sign equivalents. Semantic opacity makes these morphemes only indexable through initialization or targeting a subpart of the character as done for "peak" with the sign MOUNTAIN. As a result, it appears that the appearance of the second morpheme semantics obliges appearance of the indexation of the initial morpheme semantics due to positional salience or the tendency to represent part of the initial source name morpheme in the KSL output, which could involve a positional constraint (Beckman 2004); however, a semantically opaque initial morpheme licenses exclusive semantic indexation of the second morpheme.

	KSL Indexes source morpheme 100%	High KoTenTen Count	Swadesh	Semantically Indexed
每 매 may 'every'				
月 **월** *wel* 'moon'	o	o	o	✓
樓 누 lu 'building'				
門 **문** *mun* 'door'		o		✓
臣 신 sin 'vassal'				
龍 **용** *yong* 'dragon'	o	o		✓
梧 오 o 'Parasol tree'				
山 **산** *san* 'mountain'	o	o	o	✓

Table 7: Criteria for prototypicality

7 The Pattern of Initialization

The alternative proposal, based on ease of assimilation, hypothesizes that initialization accounts for the skipping of the initial segment. If the KSL fingerspelling for the initialization representing the first source morpheme readily assimilates to the KSL representation of the second source morpheme, then initialization happens.

Complementary distribution of KSL initialization in the dataset would support the assimilation account. Whenever an initialization appears it should license the non-representation of the initial morpheme semantics. Alternatively, ambiguous monomorphemic outputs should require indexation.

Table 1 shows that there are 8 initialized bimorphemic names; therefore, initialization does not appear to drive the non-indexation of a semantic input. Roughly half of the initializations are paired with monomorphemic names and half with polymorphemic names.

Many ambiguous monomorphemic names appear in the database; therefore, initialization is not necessarily triggered in these contexts. First considering the overall distribution of outputs from Table 1, only roughly half of the monomorphemic outputs are initialized. There are 17 initialized names in the database, of those eight are monomorphemic while nine are bimorphemic. There are 24 monomorphemic exonyms, which means that 16 are not initialized. Examples include:

雙村 쌍촌 *Ssangchon* 'twin village' → 雙 쌍 'twin'

西倉 서창 *Sechang* 'west storehouse' → 西 서'west'

松汀 송정 *Songceng* 'pine shore' → 松 송 'pine'

Output names such as 'west' or 'pine' seem ambiguous absent of the conversational context; however, even such names lack disambiguating initialization. Ji (2022) provides instances where there appears to be initialization or the addition of morphemes to disambiguate names; however,

this may reflect the disambiguation of names in the same locale. Further investigation is necessary to determine when certain ambiguous outputs are blocked and when certain ones are allowed.

With respect to the distribution of the initialized outputs, two particular Hangul characters emerged as onset indexes to greater degree relative to other input Hangul segments, ㄷ d (4/10) p=.067 and ㅂ b (3/7) p=.092. The ease of assimilation of these segments may account for their relatively high distribution (Table 3).

8 Conclusion

The evidence suggests that the appearance of an opaque initial morpheme along with a semantically salient second morpheme drives initial morpheme skipping; This semantic influence likely interacts with other social or structural constraints. Although relatively infrequent in this dataset, initialization tends to appear on KSL outputs that offer easy assimilation. Due to the small database of 100 KSL names, the hypotheses in this study require much further examination with a larger collection that can further tease out the role of semantic salience in the formation of sign language toponymic outputs.

References

Beckman, J. 2004. Positional Faithfulness. *Optimality Theory in Phonology* 1st ed., ed. J. J. McCarthy, 310–42. Malden, MA: Blackwell. https://doi.org/10.1002/9780470756171.ch16

George, J. 2022. The Lexical Shift in Japanese Sign Language (JSL) Toponyms: Accounting for the Preference of Complex over Simple Toponym Outputs. *Sign Language Studies* 23(1): 5–40. https://doi.org/10.1353/sls.2022.0011

George, J. 2025a. Structural and Social Factors in the Formation of Japanese Sign Language (JSL) Toponyms. *Sign Language Studies* 25 (3): 371–97. https://doi.org/10.1353/sls.2025.a963488.

George, J. 2025b. Japanese Sign Language (JSL) Toponym Etiology Reveals Constraints on Positional Salience in Korean Sign Language (KSL) Toponym Formation. *Japanese/Korean Linguistics* 31 (1). https://doi.org/10.5070/J7.49047.

Gwangju Metropolitan Association of the Deaf. (GMAD) 2021. "Gwangju Metropolitan City Sign Language Place Names." Gwangju Metropolitan Association of the Deaf.

Holman, E. W., S. Wichmann, C. H. Brown, V. Velupillai, A. Müller, and D. Bakker. 2008. "Explorations in Automated Language Classification." Folia Linguistica 42 (3–4). https://doi.org/10.1515/FLIN.2008.331.

Ji, Y. 2022. The Varieties of Coinage Method of the Sign Language Place Names of Gwangju Metropolitan City. *The Journal of Linguistics Science* 102: 151–90. https://doi.org/10.21296/jls.2022.9.102.151

Kelly-Jones, N., H. Hamilton, and C. B. Norris. 1981. *Signs Everywhere*. Los Alamitos: Modern Signs Press.

Kim, Y. 2019. *The Mental Representations of Hanja: Exploring Cross-Script Semantic Cohorts in Korean*. Oxford University, 215 pp.

McKee, R. L. and D. McKee. 2000. Name Signs and Identity in New Zealand Sign Language. In *Bilingualism and Identity in Deaf Communities: Sociolinguistics in Deaf Communities Series*, ed. M. Metzger, 3–40. Washington D.C.: Gallaudet University Press.

Peng, F. C. C. and D. Clouse. 1977. Place Names in Japanese Sign Language. *The Third LACUS Forum 1976*, eds. R. J. Di Pietro and E. L. Blansitt Jr., 295–308. Columbia: Hornbeam Press Inc.

Sketch Engine. KoTenTen: Korean Web Corpus 2018 (KoTenTen18). Brno: Lexical Computing CZ s.r.o., 2018. https://www.sketchengine.eu/kotenten-korean-corpus/

Stamp, R., A. Schembri, J. Fenlon, R. Rentelis, B. Woll, and K. Cormier. 2014. Lexical Variation and Change in British Sign Language. Edited by Howard Nusbaum. *PLoS ONE* 9 (4): e94053. https://doi.org/10.1371/journal.pone.0094053.

Sutrop, U. 2000. Basic Terms and Basic Vocabulary. Edited by Mati Erelt. Estonian: *Typological Studies* 4:118–45.

Sutton-Spence, R., and B. Woll. 1999. *The Linguistics of British Sign Language: An Introduction*. Cambridge, UK : New York: Cambridge University Press.