

## G-6 A contrastive study on the asymmetry of nominative and accusative case drops in Japanese and Korean

Shaoyun YU, Katsuo TAMAOKA

### 1. Introduction

#### 1.1. Variation in case marking across languages

Universally overt case marking is not a common feature in case marking languages. Among the 223 languages that exhibit a system of object case marking in Sinnemäki's (2014) survey, only 45 languages (20%) impose overt case marking on all objects, whereas the remaining 178 languages (80%) do not employ universally overt case marking. For the majority (123) of the 178 languages, the presence of object case marking is affected by either animacy or definiteness, or by a combination of the two. Japanese and Korean are among the 55 languages in which object case marking is classified as depending on other factors (Figure 1).

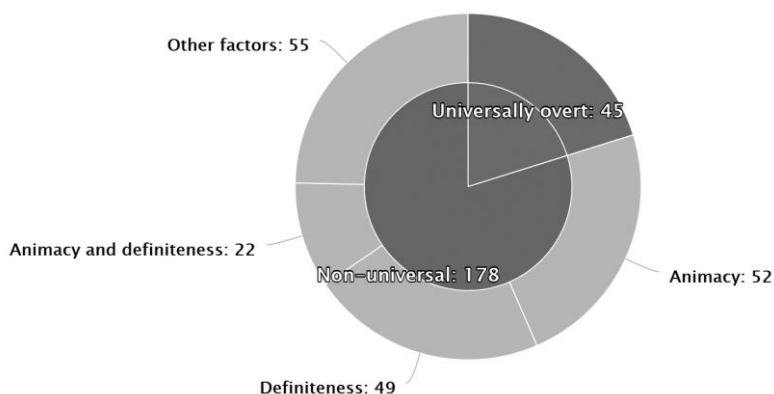


Figure 1 Object case marking patterns of 223 languages, based on Sinnemäki's (2014) data

Animacy and definiteness are two prominent factors that influence case marking in many languages. In languages using differential case marking (DCM), case marking has alternating patterns depending on a very limited number of factors, mainly the animacy and definiteness of the noun (Aissen, 2003). Aissen (2003) proposes that overt case marking in DCM languages is only applied to less typical subjects and objects, and suggests that the degree of typicality of an argument can be derived from “harmonic alignment” (Prince & Smolensky, 1993). According to harmonic alignment, the same scale (e.g., animacy, definiteness) will have mirrored ordering for subjects and objects. Take animacy for example: as shown in (1), human subjects are more typical than animate subjects, and animate subjects are more typical than inanimate subjects. On the other hand, the scale of animacy is reversed in objects: inanimate objects are more typical than animate objects, and animate objects are more typical than human objects.

- (1) Subject/Human > Subject/Animate > Subject/Inanimate  
Object/Inanimate > Object/Animate > Object/Human

In Sinhalese, for example, only the less-typical animate objects may be case-marked, while inanimate objects are not case-marked. Another example is Hebrew, in which only the less-typical definite objects are case-marked, while indefinite objects are not case-marked.

#### 1.2. Case drop in Japanese and Korean

Japanese and Korean have highly similar syntactic structures—both are verb-final SOV languages with relatively free word order, and both attach case particles to noun phrases. Moreover, the nominative and accusative case markers are frequently dropped/omitted in both colloquial Japanese and Korean (Figure 2, 3). Unlike DCM languages, case drop in Japanese and Korean is merely optional. There are not any hard rules governing case drop in Japanese and Korean, instead, case drop in these two languages is considered as “a

multi-factor phenomenon that is affected by a variety of pragmatic, semantic, syntactic, and phonological factors” (de Hoop & de Swart, 2009).

Lee (2009) argues that the major factors conditioning case marking in DCM languages (i.e., animacy, definiteness) also have effects on the phenomenon of case drop in Korean and Japanese, where they give rise to statistical preferences instead of categorical outcomes. As summarised in Figure 2, Lee’s (2009) corpus-based study shows that the drop rate of the nominative case maker in Korean, *-ka/-i*, decreases in the order: human subjects (52.7%) > animate subjects (36.0%) > inanimate subjects (28.0%), and the drop rate of the accusative case maker, *-lul/-ul*, decreases in the order: inanimate objects (74.9%) > animate (50.7%) and human (54.7%) objects. These statistical tendencies align nicely with the hierarchies in (1): the less typical an argument, the more likely it is to be overtly case-marked, or in other words, the less likely its case maker is to be dropped.

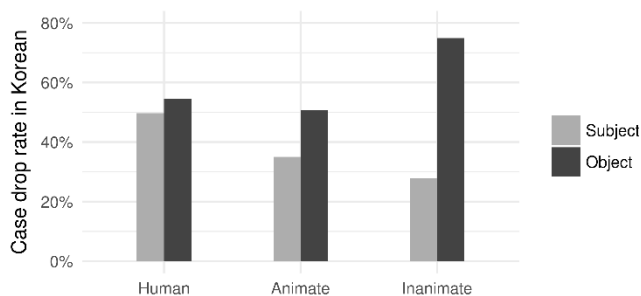


Figure 2 Plot of case drop rate in Korean, based on Lee (2009)

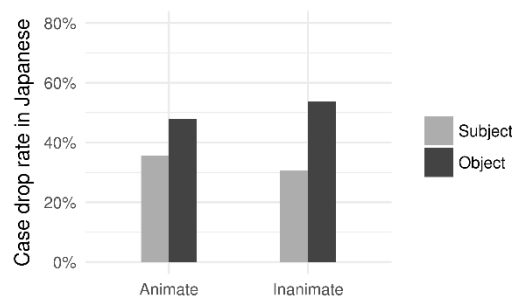


Figure 3 Plot of case drop rate in Japanese, based on Fry (2003)

The corpus-based study of Fry (2003) has reported a similar pattern to Lee (2009) in Japanese (Figure 3). The drop of the nominative case marker *-ga* is less frequent in inanimate subjects (31%) than in animate subjects (36%), and the drop of the accusative case marker *-o* is less frequent in animate objects (48%) than in inanimate objects (54%).

### 1.3. The asymmetry of nominative and accusative case drops

With the presence of either the nominative or the accusative case marker, it should be sufficient to unambiguously identify the subject and object in a transitive sentence (de Hoop & Malchukov, 2008). However, there is a nominative-accusative asymmetry in Japanese and Korean.

Ito, Tahara, & Park (1991) observed that the accusative case marker *-o* is acquired earlier than the nominative case marker *-ga* by native Japanese speakers. K. Miyahara & H. Miyahara (1973) found that the accusative case marker *-o* is frequently dropped by both adult and children native speakers of Japanese (as cited in Ito et al., 1991). Furthermore, nominative case drop is known to be more restricted than accusative case drop in Japanese, and it often leads to ungrammaticality (Saito, 1983; Fukuda, 1993).

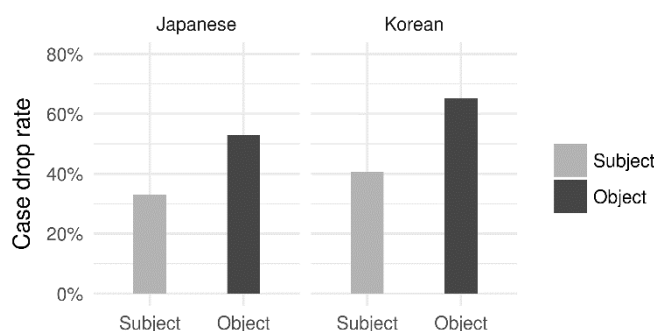


Figure 4 Comparing the overall case drop rate of the nominative and accusative case markers, based on the data from Fry (2003) and Lee (2009)

Korean exhibits a very similar nominative-accusative asymmetry: in general, the drop of the nominative case marker *-ka/-i* occurs less frequently, and is considered less acceptable than the drop of the accusative case marker *-lul/-ul* (Lee, 2012).

Indeed, by aggregating the corpus counts reported in Fry (2003) and Lee (2009), we can see that accusative case drop is more frequent than nominative case drop in both Japanese and Korean (Figure 4). In Fry's (2003) Japanese data, the accusative case marker *-o* is dropped 53% of the time and the nominative case marker is dropped 33% of the time. In Lee's (2009) Korean data, the accusative case marker *-lul/-ul* is dropped 65% of the time and the nominative case marker *-ka/-i* is dropped 41% of the time.

In Korean, case drop in the non-canonical OSV word order seems to be more restricted than case drop in the canonical SOV word order. Some linguists consider that OSV word order in Korean accommodates neither nominative nor accusative case drop (Y. Kim, 1998). Other linguists observe that only accusative case drop is allowed in OSV word order, while nominative case drop in OSV results in unacceptable sentences (Ahn & Cho, 2006).

Although previous corpus-based studies have revealed several parallelisms between Japanese and Korean in terms of the frequency of the case drop phenomenon, the perceived acceptability of case-dropped sentences is yet another important aspect to consider. A crucial point is case drop in the non-canonical OSV word order, which is difficult to study with corpus data because of the low frequency of OSV. It is reported that Japanese has an SOV:OSV ratio of 17:1 in newspaper articles (Kuno, 1978). Similarly, in colloquial Korean, T. Kim (2008) found that only 1.92% of 3,692 tokens of the sampled transitive clauses were OSV (as cited in Lee, 2016). To shed more light on the nominative-accusative asymmetry in the case drop phenomenon in Japanese and Korean, we conducted a paper-based survey that investigated the acceptability of case-dropped sentences perceived by native Japanese speakers and native Korean speakers.

## 2. Methods

The perceived acceptability of test sentences was measured with a 7-point scale acceptability judgement task. In this task, a score of 1 corresponds to the lowest acceptability and a score of 7 corresponds to the highest acceptability. The survey had a Japanese version and a Korean version, participants were tested in their native languages. Before completing the survey, participants were instructed that there were no standard answers to the judgement task and they should base their answers on their first impressions.

### 2.1. Participants

A group of native Japanese speakers and a group of native Korean speakers participated in the survey. The native Japanese speakers were college students at Nagoya University, Japan; the native Korean speakers were mainly college students at Seoul Women's University, South Korea. Participants who did not rate at least one of the two ungrammatical filler sentences with 1 (the lowest score) were excluded from the data analysis. After this exclusion, 31 native Japanese speakers (mean age: 18.8 years, *sd* = 0.7 years) and 25 native Korean speakers (mean age: 22.9 years, *sd* = 2.8 years) were included in the final data analysis.

### 2.2. Materials

Both the Japanese version and the Korean version of the survey were comprised of 64 target sentences and 4 non-target fillers. Half of the 4 non-target filler sentences were grammatical sentences and the other half were ungrammatical sentences. The sentences in the Japanese version and the Korean version were not translative equivalents.

To generate the target sentences, we first prepared a list of 8 simple transitive SOV sentences in Japanese and Korean, respectively. All the subjects of the 8 sentences were human names. In 4 of the 8 sentences, the objects were inanimate nouns, in the other 4 sentences the objects were human names. We generated the 64 target sentences by systematically manipulating the word order and case marking of the 8 original sentences. Each sentence had two word-order variants: the canonical SOV word order and the non-canonical OSV word order. For each word order, there were four case marking conditions: sentences with both the nominative and the accusative case markers (complete), sentences with the accusative case marker dropped (ACC drop), sentences with the nominative case marker dropped (NOM drop), and sentences with both case markers dropped ( $\emptyset$ ). Therefore, the target sentences were cross-classified by three factors: object animacy, word order, and case marking. It should be noticed that for sentences with human subjects and human objects, if none of the nouns is case-marked, the "OSV" variant is essentially the same with the SOV variant.

**Table 1** Examples of the target sentences

Japanese examples			Korean examples		
• misaki- $\emptyset$	tegami- <i>o</i>	kai-ta	• cinwu- $\emptyset$	khwukhi- <i>lul</i>	mek-ess-ta
Misaki- $\emptyset$	letter-ACC	write-PST	Cinwu- $\emptyset$	cookie-ACC	eat-PST-IND
<i>Misaki wrote the letter.</i>			<i>Cinwu ate the cookie.</i>		
• takashi- <i>ga</i>	masato- $\emptyset$	tasuke-ta	• senho- <i>ka</i>	mincay- $\emptyset$	mil-ess-ta
Takashi-NOM	Masato- $\emptyset$	help-PST	Senho-NOM	Mincay- $\emptyset$	push-PST-IND
<i>Takashi helped Masato.</i>			<i>Senho pushed Mincay.</i>		

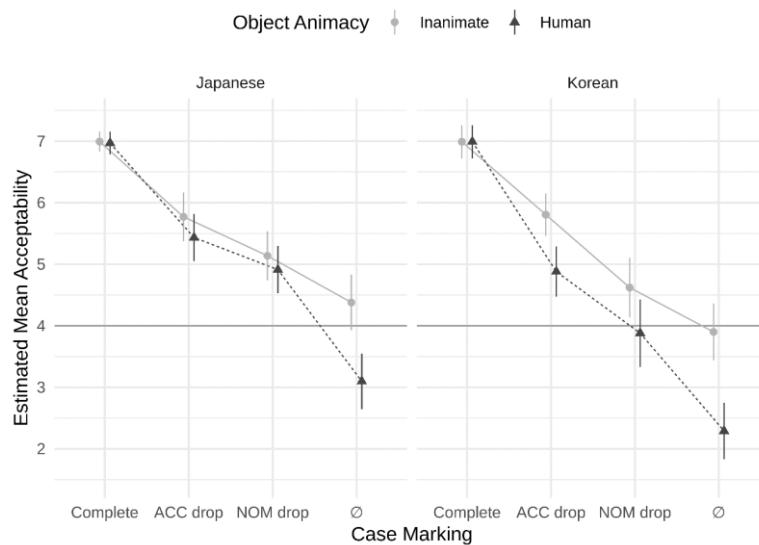
### 2.3 Data analysis

Linear mixed-effects models with random effects for participants and items (test sentences) were utilised in the analysis. The results of the Japanese survey and the Korean survey were analysed with two separate models. The structure of fixed effects was the same in the two models, which included object animacy, word order, case marking, and all interactions between the three factors. For the by-item random effects, we included the random intercept. For the by-participant random effects, we tried to include the random intercept, the random slopes for object animacy, word order, case marking, and all their interactions. While the model for the Korean data converged with this maximal specification, the model for the Japanese data required simplification. In the final model for the Japanese data, we removed the by-participant random slopes for the interaction terms.

For statistical inference, we computed the estimated mean acceptability with 95% confidence intervals (CIs) for each condition. Degrees of freedom of the estimates were determined by the Satterthwaite's approximation. For pairwise comparison of more than two levels, CIs were adjusted with Tukey's method.

## 3. Results and discussion

### 3.1. Similar patterns in SOV word order

**Figure 5** The effects of case marking and object animacy in SOV word order

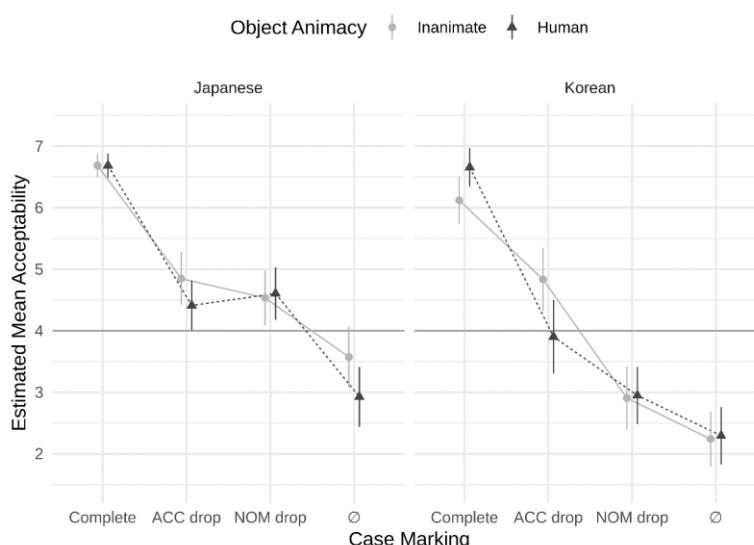
Japanese and Korean had similar patterns in SOV word order (Figure 5). The acceptability of accusative case drop was always significantly higher than that of nominative case drop in both languages. In Japanese, the difference of the acceptability between accusative and nominative case drops was 0.63, 95% CI [0.32, 0.95] for sentences with inanimate objects, and was 0.52, [0.20, 0.84] for sentences with human objects. In Korean, the difference was 1.19, [0.47, 1.90] for sentences with inanimate objects, and was 1.00, [0.21, 1.79] for sentences with human objects. While the asymmetry of nominative and accusative case drops is often dichotomized as “acceptable” or “unacceptable” in some theoretical studies, our results suggest that both nominative and accusative case drops are acceptable in Japanese and Korean SOV sentences: in most cases

their 95% CIs were above the neutral point of 4. The only exception was nominative case drop in Korean SOV sentences with human objects, in which case the acceptability was merely neutral.

Despite the overall similarities, the sizes of the estimated differences suggest that the asymmetry of nominative and accusative drops is possibly larger in Korean than in Japanese.

Our results also confirm that animacy of the object, which modulates object case marking in languages like Sinhalese, also plays a role in Japanese and Korean case drop: on the accusative case drop condition, the estimated acceptability of sentences with human objects was always lower than the acceptability of sentences with inanimate objects. However, the effects of object animacy appeared to be more pronounced in Korean. The difference of the acceptability between sentences with inanimate objects and human objects was 0.34, 95% CI [0.09, 0.58] in Japanese, which only marginally excluded zero. By contrast, the inanimate-human difference was fairly large in Korean, 0.93, [0.47, 1.38]. This result correlates nicely with previous corpus-based findings. In Lee's (2009) Korean study, the case drop rate of inanimate objects was significantly higher than that of human and animate objects (74.9% vs 53.5%,  $\chi^2=25.14$ ,  $p < 0.05$ ), by contrast, in Fry's (2003) Japanese study, the effects of object animacy did not reach statistical significance (54% vs 48%,  $\chi^2=3.68$ , *n.s.*).

### 3.2. Divergent patterns in OSV word order



**Figure 6** The effects of case marking and object animacy in OSV word order

Japanese and Korean differed dramatically in the non-canonical OSV word order (Figure 6). In Japanese, the nominative-accusative asymmetry was largely attenuated. The difference of the acceptability between accusative and nominative case drops was 0.31, 95% CI [0.00, 0.63] for sentences with inanimate objects, and was -0.19, [-0.51, 0.12] for sentences with human objects; these differences were fairly small and none was statistically significant. While the acceptability of nominative and accusative case-dropped OSV sentences was lower than their SOV counterparts, these OSV sentences were still moderately acceptable as their 95% CIs were above the neutral point of 4.

In sharp contrast to the Japanese results, Korean maintained the nominative-accusative asymmetry in OSV word order. The difference of the acceptability between accusative and nominative case drops was strikingly large for sentences with inanimate objects, 1.93, 95% CI [1.08, 2.77]. The difference for sentences with human objects also had a large point estimate, 0.95, but it did not reach statistical significance because of the wide CI, [-0.03, 1.93].

A critical difference between Japanese and Korean OSV is the poor acceptability of nominative case drop in Korean. While the estimated acceptability of nominative case drop in Japanese OSV was judged moderately acceptable, the acceptability in Korean OSV was well below the neutral point of 4, therefore was judged unacceptable. Although some studies consider that neither nominative nor accusative case drop is allowed in Korean OSV (Y. Kim, 1998), our results show that only nominative case drop is clearly ill-formed.

#### 4. Conclusion

The present study investigated the detailed acceptability of case-dropped sentences in Japanese and Korean. Our results demonstrate that both Japanese and Korean exhibit a nominative-accusative asymmetry in case-dropped SOV sentences. However, our results also revealed that the effects of case marking and object animacy might possibly vary between Japanese and Korean, despite their structural similarities. Moreover, we found Japanese and Korean differ significantly in the non-canonical OSV word order. The sharp contrast in the acceptability of nominative case drop between Japanese and Korean suggests that the phenomenon of case drop must also be regulated by certain language-specific constraints.

#### References

- Ahn, H.-D., & Cho, S. (2006). Unpronounced case markers and apparent subject-object asymmetry. *Studies in Modern Grammar*, 43, 55–80.
- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21, 435–483.
- Fry, J. (2003). *Ellipsis and wa-marking in Japanese conversation*. Routledge.
- Fukuda, M. (1993). Head government and case marker drop in Japanese. *Linguistic Inquiry*, 24(1), 168–172.
- de Hoop, H., & Malchukov, A. L. (2008). Case-marking strategies. *Linguistic Inquiry*, 39(4), 565–587.
- de Hoop, H., & de Swart, P. (2009). Cross-linguistic variation in differential subject marking. In H. de Hoop & P. de Swart (Eds.), *Differential Subject Marking* (pp. 1–16). Springer.
- Ito, T., Tahara, S., & Park, W. (1991). Acquisition of the Japanese accusative particle “o” as a patient marker. *Japanese Journal of Educational Psychology*, (39), 75–84.
- Kim, T. (2008). *Subject and object markings in conversational Korean* (Doctoral dissertation). State University of New York at Buffalo.
- Kim, Y. (1998). Overt case and covert case in Korean. *Generative Grammar Research*, 8(1), 177–237.
- Kuno, S. (1978). Japanese: A characteristic OV language. In W. P. Lehmann (Ed.), *Syntactic typology: Studies in the phenomenology of language*. (pp. 57–138). University of Texas Press.
- Lee, H. (2009). Quantitative variation in Korean case ellipsis: Implications for case theory. In H. de Hoop & P. de Swart (Eds.), *Differential Subject Marking* (pp. 41–61). Springer.
- Lee, H. (2012). Context, non-canonical word order and subject-object asymmetry in Korean case ellipsis. *Discourse and Cognition*, 19(2), 57–80.
- Lee, H. (2016). Usage probability and subject–object asymmetries in Korean case ellipsis: Experiments with subject case ellipsis. *Journal of Linguistics*, 52, 70–110.
- Miyahara, K., & Miyahara, H. (1973). Youji-ni okeru bunpou-hattatsu-no shosou [Various aspects in syntactic development of infants]. In *Proceedings of the 37th Japanese Psychological Association Meeting* (pp. 698–699).
- Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Piscataway: Rutgers University Center for Cognitive Science.
- Saito, M. (1983). Case and government in Japanese. In *WCCFL 2: Proceedings of the West Coast Conference on Formal Linguistics* (pp. 247–259).
- Sinnemäki, K. (2014). A typological perspective on Differential Object Marking. *Linguistics*, 52(2), 281–313.