‘Good-Enough’ Processing by Heritage Speakers: A Case of Korean Suffixal Passive and Morphological Causative Constructions

Gyu-Ho Shin (ghshin@uic.edu)
Department of Linguistics, University of Illinois at Chicago
601 South Morgan Street, Chicago, IL 60607 USA

Abstract

The present study investigates how heritage speakers conduct ‘good-enough’ processing at the interface of home-language proficiency, cognitive skills, and task types. For this purpose, we employ two word-order patterns of two clausal constructions in Korean (suffixal passive; morphological causative) which differ in the mapping between thematic roles and case-marking and the interpretive procedures driven by verbal morphology. We find that, while Korean heritage speakers demonstrate the same kind of acceptability-rating behaviour as monolingual Korean speakers do, their reading-time patterns are notably modulated by construction-specific properties, cognitive skills, and proficiency. This suggests a heritage speaker’s ability and willingness to conduct both parsing routes, induced by linguistic cues in a non-dominant language, which are proportional to the computational complexity involving these cues.

Keywords: ‘Good-enough’ processing; Inhibitory control; Working memory; Proficiency; Heritage speaker; Korean

Introduction

The ‘good-enough’ processing architecture (GE) assumes two processing routes: an algorithmic stream, which is a structure-based, bottom-up route, and a heuristic stream, which is a usage/experience-based, top-down route (Christianson, 2016; Ferreira, 2003). While these routes apply simultaneously to interpretation, they are distinctive concerning the trade-off between accuracy and efficiency. Algorithms yield precise computations of linguistic representations but require effortful and time-consuming processing. In contrast, heuristics allow rapid and less effortful, yet sometimes underspecified, interpretation. GE holds that the processor prioritises the heuristic route while selectively adopting the algorithmic route when required (Dwivedi, 2013; Kharkwal & Stromswold, 2014; Qian et al., 2018; Tan & Foltz, 2020). This argument is supported by various models/frameworks of sentence processing. Real-time processing places heavy demands on cognitive resources (Gibson, 1998; Lewis et al., 2006). Therefore, the processor both attempts to immediately finish interpreting input and seeks to avoid repairing misinterpretations unless urgently required (Fodor & Inoue, 1994; Piantadosi et al., 2012). Moreover, because linguistic cues are often noisy (Futrell & Levy, 2017; Gibson et al., 2019) and lossy (Christiansen & Chater, 2016), the processor favours options readily accessible from memory (e.g., *Noun–Verb–Noun* template in English; Townsend & Bever, 2001), provided that these options reasonably preserve communicative intent (Jaeger & Tily, 2010; Kleinschmidt & Jaeger, 2015). This way, the processor achieves and maintains sufficient cognitive equilibrium while minimising burdens on cognitive systems (Karimi & Ferreira, 2016; Kool et al., 2010).

Target population: Korean heritage speakers

Heritage speakers are defined as child and adult members of a linguistic minority whose home language involves limited usage experience and formal literacy education in a community and the majority language in that community is dominantly used (Montrul, 2010; Rothman, 2009). They manifest asymmetric linguistic representations influenced by various factors such as reduced home-language input, pressure on usage from the majority language, grammatical properties of a target item, and cognitive resources (Jia & Paradis, 2015; Mikhaylova, 2018; O’Grady et al., 2011; cf. Polinsky & Scontras, 2020). Previous studies have delineated distinctive attributes of heritage speakers’ morphosyntactic knowledge in comparison to L1 or L2 speakers (Felser & Arslan, 2019; Kim et al., 2009; Laleko & Polinsky, 2016; Montrul et al., 2019). Furthermore, research has elucidated the role of individual differences in modulating heritage speakers’ task performance (Bice & Kroll, 2021; Chondrogianni & Schwartz, 2020; Torres, 2023).

Our study specifically focuses on Korean heritage speakers (KHSs) residing in the United States. With over 1.9 million individuals speaking Korean as a heritage or community language in the country, this demographic constitutes the fifth-largest Asian-American subgroup (U.S. Census Bureau, 2021). Despite the increasing global interest in Korean culture and language, research within the US contexts has predominantly centred on dominant heritage speaker groups such as Hispanics (Bice & Kroll, 2021; Hur et al., 2020; Jegerski et al., 2016; López Otero et al., 2023; Torres, 2023), underscoring the urgent need for scholarly attention towards KHSs. Korean, an understudied language for GE, is a Subject–Object–Verb language that maintains verb-finality, but its case-marking system allows for relatively free word order by scrambling sentential components (Sohn, 1999). We concentrate on two clausal constructions—suffixal passive and morphological causative—which contrast with respect to alignments between thematic roles and case markers as well as interpretive procedures involving verbal morphology.

Together, we explore how KHSs engage in sentence comprehension under GE, focusing on the two parsing streams, at the intersection of proficiency (as an indicator of home-language use experience), cognitive skills (inhibitory control [IC], working memory [WM]), and task types (acceptability judgement [AJ]; self-paced reading [SPR]).

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

Two Clausal Constructions in Korean

The suffixal passive (SP) consists of two arguments, a nominative-marked theme subject and a dative-marked agent oblique, followed by a passivised verb. Passive morphology serves as a key disambiguation point for identifying a sentence's structural properties. The canonical pattern (1a) follows the theme–agent–verb ordering, but the verb can be fronted via scrambling, yielding a verb-initial pattern (1b) found in colloquial speech for afterthought clarification, information amplification, or emphasis (Sohn, 1999).

(1) Suffixal passive: 'The thief was caught by the police.'
a. Verb-final (VF)
Totwuk-i kyengchal-hanthey cap-hi-ess-ta.
thief-NOM police-DAT catch-PSV-PST-SE[1]
b. Verb-initial (VI)
Cap-hi-ess-ta totwuk-i kyengchal-hanthey.
catch-PSV-PST-SE thief-NOM police-DAT

Passive morphology in VF presents a late-arriving cue, requiring comprehenders to revise initial analyses. In Korean, a nominative-marked [+human] argument and a dative-marked [+human] argument tend to be interpreted as an agent and a recipient, respectively, which are supported by strong associations between thematic roles and case markers attested in language use (Author, xxxx; Kim & Choi, 2004; Sohn, 1999). While a plausible way of analysing (1a) prior to the verb is that the thief executes an action affecting the police, this is incongruent with the passive-voice information conveyed by verbal morphology. Thus, when encountering passive morphology, comprehenders must recalibrate the arguments' thematic roles by mapping a theme role onto the nominative-marked entity and an agent role onto the dative-marked entity, which is a cognitively demanding process (Kendeou et al., 2013; Rapp & Kendeou, 2007). In contrast, passive morphology in VI provides an early-arriving cue, facilitating accurate interpretation and mitigating thematic role misinterpretations (cf. Pozzan & Trueswell, 2015).

The morphological causative (MC) consists of three arguments: a subject (causer), an indirect object (causee), and a direct object (theme), as in (2a). The verb carries one of the seven allomorphic variants of verbal suffixes. The verb can move to the sentence-initial position as in (2b). The interpretation of the arguments' thematic roles hinges upon causative morphology, but this process does not invoke substantial challenges to the extent that passive morphology does. To illustrate, in (2a), the nominative-marked [+human] argument *Mia-ka* is understood as a causer (as an extension of an agent, sharing the concept of a volitional actor). The dative-marked [+human] argument *Pola-eykey* is understood as a causee (as an extension of a recipient); the dative marker ensures these extensions by sharing the same semantic component—GOAL (Sohn, 1999). Causative morphology does not invite the same kind of recalibration of the mapping between thematic roles and case markers as that needed in passive morphology. Therefore, the degree of cognitive burdens that verbal morphology poses to processing the morphological causative is not enormous.

(2) Morphological causative: 'Mia made Pola eat food.'
a. VF
Mia-ka Pola-eykey umsik-ul mek-i-ess-ta.
Mia-NOM Pola-DAT food-ACC eat-CST-PST-SE
b. VI
Mek-i-ess-ta Mia-ka Pola-eykey umsik-ul.
eat-CST-PST-SE Mia-NOM Pola-DAT food-ACC

Methods[2]

We recruited 40 KHSs (M_{age} = 24.0, SD = 5.2) who were born in the USA, were raised by Korean-speaking parents, and had resided in America for most of their lives (length of stay in the USA: M = 21.9, SD = 6.2). They use English more frequently than Korean in daily life (English: M = 92.5, SD = 9.5; Korean: M = 37.1, SD = 27.3; score out of 100) and adopt Korean more often with family than colleagues (family: M = 4.98, SD = 1.25; friends: M = 3.45, SD = 1.38; colleagues: M = 3.25, SD = 1.63; score out of 6 [1 = English only; 6 = Korean only]). They expressed greater confidence in their proficiency in listening to and speaking Korean (M = 4.03, SD = 0.92 [0 = not good; 5 = very good]) compared to their skills in reading and writing Korean (M = 3.05, SD = 1.20 [0 = not good; 5 = very good]), also confirmed by a one-sample t-test: $t(78) = 4.085$, $p < .001$. Nevertheless, they expressed dissatisfaction with their ability to speak Korean (M = 2.83, SD = 1.39 [0 = not satisfied; 5 = very satisfied]) and perceived their command of Korean as falling short of target-like use (M = 2.05, SD = 1.66 [0 = fully disagree; 5 = fully agree]). All the KHSs in this study learnt Korean primarily from their parents, supplemented by additional exposure through three major channels: educational institutions such as language schools, universities, and academies (80%), online resources (70%), and social interactions with friends and peers (70%). We also recruited 32 native speakers of Korean (NSK; M_{age} = 25.7, SD = 4.3) as a control group.

Participants joined a *Zoom* meeting and completed the tasks individually on web-based platforms: proficiency (a *JavaScript*-based platform), cognitive task (*PsyToolkit*; Stoet, 2010, 2017), SPRT (*PCIbex*; Zehr & Schwarz, 2018), AJ (*Qualtrics*), and background survey (*Google Forms*). For the stability of testing environments, mobile devices were prohibited. Their participation in MC was delayed one week after their participation in SP; during their second participation, they did only AJ and SPR.

Proficiency in Korean was assessed using the Korean C-test (Lee-Ellis 2009), which evaluates comprehension through five passages with syllable-unit blanks. We chose the first four passages for efficiency, following the original study's recommendation. Each blank represented one point,

[1] Abbreviations: ACC = accusative case marker; CST = causative suffix; DAT = dative marker; NOM = nominative case marker; PST = past tense marker; PSV = passive suffix; SE = sentence ender; V = verb.

[2] See this repository for the data/code and detailed model outcomes of this study.

with a maximum score of 188. The proficiency scores of participants (M = 127.3, SD = 25.8) exceeded those of L2 learners in Lee et al. (2023), indicating that KHS possessed commendable literacy and reading skills in home language.

We measured participants' WM via a digit-span task (Miller, 1956) considering its popularity in the field, simplicity of implementation and interpretation, and superiority to other measurement types (Baddeley et al., 1998; Jones & Macken, 2015; Schofield & Ashman, 1986). Participants were exposed to a sequence of two digits. A longer sequence was presented if they succeeded in repeating the sequence and until they failed to repeat it correctly. The longest sequence that they retrieved correctly was considered their digit span. To ensure this measure's reliability, sequence length was increased after recall of the same length twice. We also measured participants' IC by slightly adapting the Flanker task used by Eriksen and Eriksen (1974). Participants were presented with five letters and instructed to respond to the one in the middle by pressing 'A' on the keyboard when they saw 'X' or 'C' and pressing 'L' upon perceiving 'V' or 'B'. We counted the total number of correct responses out of 50 trials, whether under congruent or incongruent conditions, and excluded excessively slow responses (i.e., reaction time > 3000 ms). Each task lasted for around five minutes.

[SPR] We created 32 test sentences (8 VF & 8 VI per construction), each comprising a carrier phrase (e.g., *Nay-ka tul-ess-nuntey*, 'I heard that' [SP]; *ceki N* 'N over there' [MC]), followed by the critical structure and a temporal adverbial phrase (Table 1). For agent/theme nominals, we used human names often attested in daily life; all the verbs (with sufficient usage frequency) were expressed in the past tense, and no overlap occurred in verb use across the construction types. When creating the MC sentences, to make critical and spill-over regions as comparable as possible across the two constructions, we omitted an accusative case marker of the direct object and topicalised it by moving it to the sentence-initial position; the target frame contained a nominative-marked NP, a dative-marked NP, and a verb.[3] Each item was presented in six regions (Rs), with R2–R4 as the critical regions and R5 as a region for accommodating the spill-over effects induced by a task-specific button-press strategy. The test sentences were interspersed with 64 fillers of various structures and complexities.

Table 1. Scheme of test sentences (SPR).

Cx	Cond	R1	R2	R3	R4	R5	R6
SP	VF	*I heard*	N-NOM	N-DAT	V-PSV	*yester-day*	*night*
	VI	*that*	V-PSV	N-NOM	N-DAT		
MC	VF	*N over*	N-NOM	N-DAT	V-CST		
	VI	*there*	V-CST	N-NOM	N-DAT		

SPR was run under a non-cumulative moving-window paradigm (Just et al., 1982), with each target sentence appearing at the centre of the screen on a region-by-region basis. In the beginning of each trial, participants saw a series of dashes on-screen, and each press of a spacebar revealed words in each region while concealing preceding words. Following each sentence, a simple comprehension question appeared to direct participants' attention to the task. Participants responded by clicking on one of two choices, and upon the choice of an erroneous answer, a 'wrong choice' feedback appeared on-screen. Each question involved simple facts regarding the sentence being read (e.g., what the sentence was about, what action was done), in contrast to previous studies wherein questions asked about an agent or a theme and answers served as reflections of comprehenders' misunderstanding (e.g., Ferreira, 2003). We used participants' responses only as an attention check (cf. Dwivedi, 2013). Prior to the experiment, they received written instructions and worked through three practice items for familiarisation with the procedures. The task took approximately 20 minutes.

Data were first trimmed by excluding the RT datapoints of all the regions in a sentence upon failure in the comprehension check for that sentence (data loss: 1.06% [SP]; 1.07% [MC]) and by excluding outliers per region through a 3SD cut-off point (collapsing over item and participant; data loss: 2.94% [SP]; 3.13% [MC]). We then log-transformed the pruned data for normalisation and residualised them to adjust for variability in word length and individual reading speed (Baayen & Milin, 2010). The pre-processed data were fitted to linear mixed-effect modelling for each critical and spill-over region per construction, with *Group* and *Condition* as fixed effects (centred around the mean and deviation-coded) and with *Participant* and *Word* as random effects using *lme4* (Bates et al., 2015) in *R* (R Core Team, 2023). For KHS-internal models, the three factors (*Digit, Flanker, Proficiency*) were treated as continuous variables and included as fixed effects; each model consisted of only two fixed effects (Condition and one of these factors), resulting in three sub-models per region. The other specifications were the same as those in the global model.

[AJ] The sentences for AJ were created by clipping the main regions of test sentences used in SPR (SP_VF/VI: R2 + R3 + R4; MC_VF: R2 + R3 + N-ACC + R4; MC_VI: R2 + R3 + R4 + N-ACC). The clausal composition of the test sentences across the two construction types differed because of the topicalised theme object in MC. We acknowledge that it could have been a confound in precisely revealing task effects through this contrast.

Only one sentence appeared on the screen per trial. Participants rated the acceptability of each sentence with a 6-point Likert scale (unacceptable: 0; acceptable: 5), responding immediately upon encountering the sentence but

[3] Compared to the general conversation initiator at R1 in SP, the case-less noun at R1 in MC may provide a better background for a comprehender to handle the fronted verb, possibly reducing surprisal/disequilibrium when processing R2 for VI to some degree. We concede that the different compositions of R1 across the two construction types, albeit unavoidable, could have confounded the interpretation of the results.

without sacrificing the accuracy and faithfulness of/confidence in their response. Once participants clicked on the scale and moved on to the next sentence, they were prohibited from revising their previous evaluation. This task was untimed and took approximately 15 minutes.

Data were trimmed by excluding the individual values with response times below 1,000 ms or above 10,000 ms (data loss: 5.55% [SP]; 6.25% [MC]). We then Z-transformed the pruned data for normalisation and proceeded to the same kind of linear mixed-effect modelling per construction, with *Group* and *Condition* as fixed effects (centred around the mean and deviation-coded) and with *Participant* and *Sentence* as random effects (Bates et al., 2015) in *R* (R Core Team, 2023). For KHS-internal models, the three factors (*Digit, Flanker, Proficiency*) were treated as continuous variables and included as fixed effects; each model consisted of only two fixed effects (*Condition* and one of these factors), resulting in four sub-models per region. The other specifications were the same as those in the global model.

Predictions

[Suffixal passive] In SPR, we predict that KHS should spend more time reading sentences than NSK given the global difficulty in the real-time processing of non-dominant languages (Grüter & Rohde, 2021; Hopp, 2014) and the reduced degree/richness of home-language exposure (Jia & Paradis, 2015; Unsworth, 2013). We also expect KHS's increased RT spent at/after the verb due to the interpretive procedures involving passive morphology. In addition, we anticipate comparable RTs between the verb regions of the two conditions due to the competing dynamics of heightened surprisal/disequilibrium associated with verb-initiality at R2 in VI versus the increased cognitive load incurred by the necessary revision process at R4 in VF. We anticipate differences in how the three factors (WM, IC, proficiency) influence KH''s RT patterns. The sentence-initial verb generates more surprisal, but IC helps manage this, leading to decreased RTs with higher IC capacities. WM aids in efficiently revising the initial interpretation required by the sentence-final verb, resulting in reduced RTs with larger WM capacities. Proficiency enhances overall processing efficiency, leading to decreased RTs across both patterns.

In AJ, KHS is expected to prefer the VF/canonical pattern over the VI/non-canonical counterpart, aligning with previous findings on heritage speakers' performance (Chondrogianni & Schwartz, 2020; Kim et al., 2018). This preference arises from the typical and frequent sentence structure in Korean (subject-first + predicate-final) and the contextual effects of scrambling. As proficiency and cognitive skills increase, the rating gap between the patterns is anticipated to widen, indicating greater recognition of the unnaturalness of scrambled sentences.

[Morphological causative] We expect KHS to exhibit similar acceptability-rating trends, with difficulties in real-time processing, as in Study 1. However, we anticipate that KHS would demonstrate faster RTs in verb-related regions in SPR due to the less demanding nature of interpretive procedures involving causative morphology. Additionally, this nature may lead to KHS's heightened reaction to verb-initiality over verb-finality, potentially resulting in longer RTs at R2 in VI compared to at R4 in VF. We expect the influence of the three factors (WM, IC, proficiency) on KHS's RT patterns to vary, similar to the case of SP. Specifically, given the less radical realignment between thematic roles and case markers driven by causative morphology compared to passive morphology, we anticipate that IC would more effectively manage the surprisal/disequilibrium generated in VI, resulting in decreased RTs when processing the fronted verb proportional to IC capacities.

Results

Cognitive skills

For the digit-span task, the mean score of KHS was 6.6 (*SD* = 1.3). When compared to that of NSK (*M* = 7.8, *SD* = 1.1), the two groups differed (independent-sample t-test: $t(70) = -4.062$, $p < .001$). For the Flanker task, the mean score of KHS was 42.3 (*SD* = 6.3). When compared to that of NSK (*M* = 38.9, *SD* = 11.8), the two groups did not differ (independent-sample t-test: $t(70) = 1.552$, $p = .125$).

Self-paced Reading

In SP, the global model ($\alpha = .05$) revealed main effects of *Group* at all the regions of interest and *Condition* at R2 and R5. Additional analyses ($\alpha = .025$) showed no difference at each region for NSK but significant differences at R2 ($p = .002$) and R5 ($p = .002$) for KHS. These indicate that, given the overall by-group difference (R2 to R4: KHS > NSK; R5: KHS < NSK), KHS demonstrated notable by-condition variances (R2: VF < VI; R5: VF > VI). A verb-region model (fixed effects: *Group, Condition*; random effect: *Participant*; $\alpha = .025$) revealed only a main effect of *Group* ($p < .0005$), indicating that each group spent comparable RTs across the two verb regions.

In MC, the global model ($\alpha = .05$) revealed main effects of *Condition* and *Group* at R2, a main effect of *Group* at R3, and an interaction effect between *Condition* and *Group* at R5. Post-hoc analyses ($\alpha = .025$) revealed no RT difference at each region for NSK and a significant RT difference only at R5 for KHS. These indicate that, given the overall RT difference by group (R2 & R3: KHS > NSK), KHS demonstrated notable by-condition RT difference at R5 (VF > VI). This is partially consistent with the case of SP, except that by KHS at R2 and that by the two groups at R4. Notably, a verb-region model (fixed effects: *Group, Condition*; random effect: *Word*; $\alpha = .025$) revealed main effects of *Group* ($p < .0005$) and *Condition* ($p = .005$) and interaction ($p = .016$), with a significant between-condition difference only for KHS ($p < .0005$). In addition, KHS spent less time reading R4 in VF of Study 2 compared to Study 1 ($p = .004$; $\alpha = .0125$). These findings indicate a substantial difference in the RTs that KHS allocated to the verb regions across the two conditions.

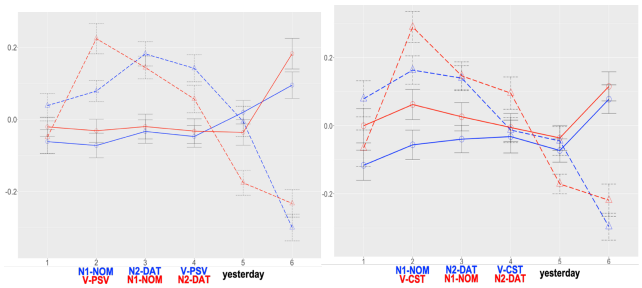

Figure 1: Results (SPR). Left: SP; Right: MC. X-axis: region; Y-axis: residual RT. Blue: VF; Red: VI. Solid line: NSK; dashed line: KHS. Error bars: 95% CI.

KHS-internal models (SP) (α = .025) At R2, we found an interaction effect between *Condition* and *Flanker*, and post-hoc analyses (α = .0125) uncovered marginal significance in VF and insignificance in VI. This trend was supported by the correlation analysis in which the association between the Flanker scores and the RTs was meaningful only in VF (r = –0.301, p = .009). These indicate that KHS spent less time reading R2 in VF as their IC capacities expanded. At R4, we found a main effect of *Proficiency*; additional analyses (α = .0125) yielded insignificance in VF and significance in VI. This trend was supported by the correlation analysis, with the association between the proficiency scores and the RTs being meaningful only in VI (r = –0.361, p = .001). These indicate that, given the broad impact of proficiency on RTs at this region, KHS spent less time particularly in VI with increasing proficiency. At R5, we found a marginal interaction effect between *Condition* and *Digit*, and additional analyses (α = .0125) yielded insignificance in both conditions. However, a meaningful relationship existed between the digit-span scores and the RTs in VF, as shown by the correlation analysis (r = 0.231, p = .047). These indicate that KHS spent more time reading R5 in VF with larger (albeit weak) WM capacities. Neither of verb-region models (fixed effects: *Condition* and one of the following factors [*Digit, Flanker, Proficiency*]; random effect: *Participant* for model convergence; α = .025) revealed significant main or interaction effects (all ps > .1).

KHS-internal models (MC) (α = .025) At R4, we found a main effect of *Proficiency*; additional analyses (α = .0125) yielded insignificance in VF and significance for VI. This trend was supported by the correlation analysis: the association between the proficiency scores and the RTs was meaningful only in VI (r = –0.334, p = .003), indicating that KHS spent less time reading R4 in VI with increasing proficiency. At R5, we found a marginal interaction effect between *Condition* and *Digit* and an interaction effect between *Condition* and *Flanker*. Post-hoc analyses (α = .0125) yielded insignificance in both conditions, but meaningful relationships were found between the scores of the two tasks and the RTs in VF, as shown by the correlation analysis (r = 0.230, p = .048 for *Digit*; r = 0.237, p = .039 for *Flanker*). These indicate that KHS spent more time reading R5 in VF with expanding (albeit weak) IC or WM capacities. Neither of verb-region models (fixed effects: *Condition* and one of the following factors [*Digit, Flanker, Proficiency*]; random effect: *Word* for model convergence; α = .025) revealed significant main or interaction effects (all ps > .1).

Acceptability Judgement

Both groups preferred VF over VI, but the by-condition gap was larger for KHS than NSK. The global model (α = .05) revealed a main effect of *Condition* and a *Condition–Group* interaction in SP and a main effect of *Condition* in MC. Post-hoc analysis (α = .025) per construction type yielded insignificance for all the between-condition comparisons within each group, indicating uniformity in the two groups' preference for VF in both constructions.

KHS-internal models (α = .025) In SP, the models revealed interaction effects between *Condition* and each factor (*Digit, Flanker, Proficiency*), indicating that KHS evaluated VI as less acceptable with increasing WM capacities, IC capacities, or proficiency. In MC, the models revealed an interaction effect only in the KHS–Flanker model. This indicates that the Flanker scores modulated their ratings, driving KHS to evaluate VI as less acceptable with increasing IC capacities.

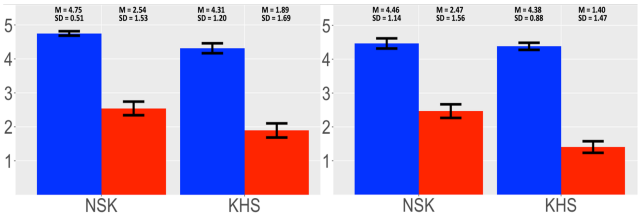

Figure 2: Results (AJ). Left: SP; Right: MC. X-axis: Group; Y-axis: acceptability (raw). Blue: VF; Red: VI. M: mean; SD: standard deviation. Error bars: 95% CI.

Discussion and Conclusion

While NSK exhibits a strong preference for VF over VI in AJ, no processing advantage is observed in SPR. These suggest a heuristic-before-algorithm strategy for sentence processing (Dwivedi, 2013; Kharkwal & Stromswold, 2014). KHS's performance sheds light on how GE interacts with various factors during comprehension. Despite showing comparable acceptability-rating behaviour to NSK in AJ, KHS demonstrated prolonged RTs at critical regions in both construction types in SPR. These findings align with prior research indicating the inherent challenges in real-time processing of non-dominant languages (Grüter & Rohde, 2021; Pozzan & Trueswell, 2016).

Suffixal Passive (SP)

In SPR, the fronted verb in VI incurred greater processing cost compared to the nominative-marked subject in VF, whereas the post-verbal region in VI incurred reduced processing cost compared to that region in VF. Interestingly, KHS demonstrated similar RTs when reading the verb region in both conditions, which suggests a limited role for early-appearing verbal morphology cues in processing verb-initial passive sentences. This highlights a stronger influence of heuristic parsing over algorithmic parsing: the expected benefit from fronted verbal morphology may not outweigh

the advantage of word-order canonicity, readily accessible from memory and potentially alleviating interpretive challenges posed by passive morphology. Multiple factors, including heritage speakers' susceptibility to home-language morphosyntax (Chondrogianni & Schwartz, 2020; Kim et al., 2018), inflexibility in dealing with scrambling due to dominance of English (Namboodiripad et al., 2018), and limited usage experience of target linguistic knowledge in the home language (Hur et al., 2020; López Otero et al., 2023), may be ascribed to these findings.

Notably, KHS's performance was modulated by cognitive skills and proficiency. This interplay was more complex in SPR than AJ: KHS's acceptability ratings were proportionate to their scores on the three measures (digit span, Flanker, proficiency), but the contributions of these measures to their RT patterns varied at different regions and conditions. KHS spent less time reading R2 (verb in VI vs. noun in VF) in VF with expanding IC capacities, but this trend did not emerge in VI. This difference is ascribed to an increased degree of interpretive challenge involving verb-initiality. That is, the fronted verb—manifesting atypical word order and inviting (re)calibrations of the mapping between thematic roles and case markers early on—may have substantially cancelled out processing support from IC when KHS handled the early-appearing verb/morphology cue.

KHS spent less time reading R4 in VI with increasing proficiency. This suggests two interpretations. First, as proficiency increased, KHS demonstrated greater proficiency in handling passive morphology (and associated algorithmic parsing) in VF. Second, it implies improved efficiency in processing a dative-marked agent in VI with increasing proficiency. However, the expected overall influence of proficiency across the entire structure did not emerge, suggesting the selective contribution of general proficiency in a non-dominant language to sentence processing contingent upon task types (cf. Robert, 2012).

KHS spent more time reading R5 (spill-over involving clausal integration for complete interpretation) in VF as their WM capacities expanded although the effect was weak. This finding implies KHS's increased capacity for, and commitment to, the integration procedures involving the canonical word-order condition proportionate to their WM skills. Such capacity and commitment, in turn, enable a comprehender to reserve more space for coping with previous and current inputs at this region.

Morphological Causative (MC)

KHS demonstrated a preference for the verb-final condition over the verb-initial condition in AJ, with significant RT differences at verb-related regions. The insignificant by-condition difference at R2 may suggest a potential benefit from early-arriving cues. However, caution is warranted as this appears to stem from notably increased RTs in VF compared to VI, for reasons yet unclear. The longer RTs for KHS when reading R2 in VI compared to R4 in VF suggests the larger role of heuristic parsing than algorithmic parsing in sentence processing. Furthermore, the reduction in RTs at R4 in VF of Study 2, despite the cognitive demands of clausal integration (indicated by the RT gap at R5), suggests that KHS may have relied on interpretive procedures involving causative morphology, assumed to be less demanding than those involving passive morphology, to some extent.

Pertaining to proficiency and cognitive skills, only the Flanker scores influenced the KHS's acceptability ratings, diverging from the findings in SP. This suggests that construction-specific properties (e.g., mapping of thematic roles and case markers, interpretive procedures driven by verbal morphology) selectively adjust the activation of these factors to varying degrees in this construction. In SPR, while the proficiency and digit-span scores influenced KHS's RT patterns similarly to SP, the Flanker scores led to increased RTs at R5 in VF (albeit weakly). The role of proficiency found here aligned with that in SP, supporting the notion that general proficiency in a non-dominant language selectively contributes to sentence processing in that language.

The observed asymmetry in IC may arise from differences in the properties of the two construction types examined in this study. Relative to the interpretive procedures involving SP, those involving MC are less demanding. This qualitative difference in the algorithmic stream applied to each construction type may have allowed KHS to allocate more cognitive resources to clausal integration at R5, resulting in increased RTs at this region. This pattern is also consistent with KHS spending more time at R5 in VF relative to the digit-span scores in the case of SP. In this respect, the absence of Flanker effects at R2 in VI in MC, as well as at R2 in VI and at R5 in VF in SP, point to the same mechanism underlying non-dominant-language mind: its capacity and readiness to engage in (algorithmic) parsing induced by linguistic cues, proportionate to the computational complexity involving these cues. To illustrate, scrambled word order presents greater computational challenge compared to canonical word order due to its infrequent usage and potential contextual/discoursal effects on interpretation. This may have mitigated the presumed early-arriving-cue advantage in VI of both constructions. Passive morphology imposes greater interpretive demands than causative morphology, potentially limiting the non-dominant-language mind's ability to filter out irrelevant information via IC.

Concluding Remarks

By examining two Korean construction types that differ in two parsing streams, we have revealed the interface between GE as a sentence-processing architecture and various factors surrounding heritage speakers such as home-language proficiency, cognitive skills, and task types. The processor prioritises efficiency by minimising cognitive demand and processing effort, strategically employing heuristics and algorithms in response to linguistic cues. Concurrently, various (non-)linguistic factors dynamically affect sentence processing, constructing noisy representations of non-dominant-language knowledge (Futrell & Gibson, 2017; Tachihara & Goldberg, 2020). Our experimental settings offer insights into this aspect, thereby advancing the understanding of a learner's mind for underrepresented languages and populations in the field.

References

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28. https://doi.org/10.21500/20112084.807

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173. https://doi.org/10.1037/0033-295X.105.1.158

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. http://hdl.handle.net/10.18637/jss.v067.i01

Bice, K., & Kroll, J. F. (2021). Grammatical processing in two languages: How individual differences in language experience and cognitive abilities shape comprehension in heritage bilinguals. *Journal of Neurolinguistics*, *58*, 100963. https://doi.org/10.1016/j.jneuroling.2020.100963

Chondrogianni, V., & Schwartz, R. G. (2020). Case marking and word order in Greek heritage children. *Journal of Child Language*, *47*(4), 766–795. https://doi.org/10.1017/S0305000919000849

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences, 39*. https://doi.org/10.1017/S0140525X1500031X

Christianson, K. (2016). When language comprehension goes wrong for the right reasons: Good-enough, underspecified, or shallow language processing. *Quarterly Journal of Experimental Psychology, 69*(5), 817–828. https://doi.org/10.1080/17470218.2015.1134603

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

Felser, C., & Arslan, S. (2019). Inappropriate choice of definites in Turkish heritage speakers of German. *Heritage Language Journal*, *16*(1), 22–43. https://doi.org/10.46538/hlj.16.1.2

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–203. https://doi.org/10.1016/S0010-0285(03)00005-7

Fodor, J. D., & Inoue, A. (1994). The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research*, *23*, 407–434. https://doi.org/10.1007/BF02143947

Futrell, R., & Gibson, E. (2017). L2 processing as noisy channel language comprehension. *Bilingualism: Language and Cognition, 20*(4), 683–684. https://doi.org/10.1017/S1366728916001061

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. Cognition, 68(1), 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1

Gibson, E., Futrell, R., Piantadosi, S. P., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences, 23*(5), 389–407. https://doi.org/10.1016/j.tics.2019.02.003

Grüter., T., & Rohde, H. (2021). Limits on expectation-based processing: Use of grammatical aspect for co-reference in L2. *Applied Psycholinguistics, 42*, 51–75. https://doi.org/10.1017/S0142716420000582

Hopp, H. (2014). Working memory effects in the L2 processing of ambiguous relative clauses. *Language Acquisition*, *21*(3), 250–278. https://doi.org/10.1080/10489223.2014.892943

Hur, E., Lopez Otero, J. C., & Sanchez, L. (2020). Gender agreement and assignment in Spanish heritage speakers: Does frequency matter?. *Languages*, *5*(4), 48. https://doi.org/10.3390/languages5040048

Jaeger, T. F., & Tily, H. (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(3), 323–335. https://doi.org/10.1002/wcs.126

Jegerski, J., Keating, G. D., & VanPatten, B. (2016). On-line relative clause attachment strategy in heritage speakers of Spanish. *International Journal of Bilingualism*, *20*(3), 254–268. https://doi.org/10.1177/1367006914552288

Jia, R., & Paradis, J. (2015). The use of referring expressions in narratives by Mandarin heritage language children and the role of language environment factors in predicting individual differences. *Bilingualism: Language and Cognition*, *18*(4), 737–752. https://doi.org/10.1017/S1366728914000728

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes and in reading comprehension. *Journal of Experimental Psychology: General, 3*(2), 228–238. https://doi.org/10.1037/0096-3445.111.2.228

Karimi, H., & Ferreira, F. (2016). Good-enough linguistic representations and online cognitive equilibrium in language processing. *Quarterly Journal of Experimental Psychology, 69*(5), 1013–1040. https://doi.org/10.1080/17470218.2015.1053951

Kendeou, P., Smith, E. R., & O'Brien, E. J. (2013). Updating during reading comprehension: why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 854–865. https://doi.org/10.1037/a0029468

Kharkwal, G., & Stromswold, K. (2014). Good-enough language processing: Evidence from sentence-video matching. *Journal of Psycholinguistic Research*, *43*(1), 27–43.

Kim, J-B., & Choi, I. (2004). The Korean case system: A unified, constraint-based approach. *Language Research, 40*, 885–921.

Kim, K., O'Grady, W., & Schwartz, B. D. (2018). Case in heritage Korean. *Linguistic Approaches to Bilingualism*, *8*(2), 252–282. https://doi.org/10.1075/lab.16001.kim

Kim, S. Y., Sung, J. E., & Yim, D. (2017). Sentence comprehension ability and working memory capacity as a function of syntactic structure and canonicity in 5-and 6-year-old children. *Communication Sciences & Disorders, 22*(4), 643–656. https://doi.org/10.12963/csd.17420

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*(2), 148–203. https://doi.org/10.1037/a0038695

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139*(4), 665–682. https://doi.org/10.1037/a0020198

Lee, C., Shin, G-H., & Jung, B. K. (2023). How 'good-enough' is second language comprehension? Morphological causative and suffixal passive constructions in Korean. *Applied Linguistics Review*. https://doi.org/10.1515/applirev-2022-0152

Lee-Ellis, S. (2009), The development and validation of a Korean C-test using Rasch analysis. *Language Testing, 26*(2), 245–274. https://doi.org/10.1177/0265532208101007

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454. https://doi.org/10.1016/j.tics.2006.08.007

Lim, J. H., & Christianson, K. (2015). Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation. *Applied Psycholinguistics, 36*(6), 1283–1315. https://doi.org/10.1017/S0142716414000290

López Otero, J. C., Hur, E., & Goldin, M. (2023). Syntactic optionality in heritage Spanish: How patterns of exposure and use affect clitic climbing. *International Journal of Bilingualism*, 13670069231170691. https://doi.org/10.1177/13670069231170691

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158

Montrul, S. (2010). How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics, 31*(1), 167–207. https://doi.org/10.1017/S014271640999021X

Namboodiripad, S., Kim, D., & Kim, G. (2018). English dominant Korean speakers show reduced flexibility in constituent order. In Daniel Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Sankhagowit Amara & Michael Tabatowski (eds.), *Proceedings of the Fifty-Third Annual Meeting of the Chicago Linguistic Society* (pp. 247–260). Chicago: Chicago Linguistic Society.

O'Grady, W., Kwak, H. Y., Lee, O. S., & Lee, M. (2011). An emergentist perspective on heritage language acquisition. *Studies in Second Language Acquisition, 33*(2), 223–245. https://doi.org/10.1017/S0272263110000744

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition, 122*, 280–291. https://doi.org/10.1016/j.cognition.2011.10.004

Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition*, *23*(1), 4–20. https://doi.org/10.1017/S1366728919000245

Pozzan, L., & Trueswell, J. C. (2015). Revise and resubmit: How real-time parsing limitations influence grammar acquisition. *Cognitive Psychology*, *80*, 73–108.

Qian, Z., Garnsey, S., & Christianson, K. (2018). A comparison of online and offline measures of good-enough processing in garden-path sentences. *Language, Cognition and Neuroscience, 33*(2), 227–254. https://doi.org/10.1080/23273798.2017.1379606

R Core Team. (2023). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, *35*(8), 2019–2032. https://doi.org/10.3758/BF03192934

Roberts, L. (2012). Individual differences in second language sentence processing. *Language Learning*, *62*, 172–188. https://doi.org/10.1111/j.1467-9922.2012.00711.x

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, *13*(2), 155–163. https://doi.org/10.1177/1367006909339814

Schofield, N. J., & Ashman, A. F. (1986). The relationship between digit span and cognitive processing across ability groups. *Intelligence, 10*(1), 59–73. https://doi.org/10.1016/0160-2896(86)90027-9

Sohn, H. M. (1999). *The Korean language*. Cambridge: Cambridge University Press.

Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Tachihara, K., & Goldberg, A. E. (2020). Reduced competition effects and noisier representations in a second language. *Language Learning, 70*(1), 219–265. https://doi.org/10.1111/lang.12375

Tan, M., & Foltz, A. (2020). Task sensitivity in L2 English speakers' syntactic processing: Evidence for good-enough processing in self-paced reading. *Frontiers in Psychology*, 11. https://doi.org/10.3389/fpsyg.2020.575847

Torres, J. (2023). Exploring working memory and language dominance in heritage bilinguals' writing processes. *Studies in Second Language Acquisition*, *45*(3), 710–736. https://doi.org/10.1017/S027226312300013X

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, *16*(1), 86–110. https://doi.org/10.1017/S1366728912000284

U.S. Census Bureau. (2021). Asian American and Pacific Islander heritage month: May 2021. Retrieved from https://www.census.gov/newsroom/facts-for-features/2021/asian-american-pacific-islander.html (accessed 17 April 2024)

Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments* (IBEX). https://doi.org/10.17605/OSF.IO/MD832