How 'Good-Enough' is L2 Sentence Comprehension? Evidence from Suffixal Passive Construction in Korean

Chanyoung Lee (cy.lee@yonsei.ac.kr)
Department of Korean Language and Literature, Yonsei University
50 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Gyu-Ho Shin (gyuho.shin@upol.cz)
Department of Asian Studies, Palacký University Olomouc
tř. Svobody 26, Olomouc, Czech Republic

Boo Kyung Jung (boj11@pitt.edu)
Department of East Asian Languages and Literatures, University of Pittsburgh
4200 Fifth Ave, Pittsburgh, PA, United States of America

Abstract

This study investigates how L2 learners achieve the 'good-enough' comprehension in Korean. We focus on a suffixal passive construction, given the scarcity of this construction in the L2 textbook input. Results from acceptability judgement and self-paced reading tasks suggest two aspects of L2 comprehension. First, L1 and L2 comprehension do not qualitatively differ regarding 'good-enough' processing: the L2 processor utilises both heuristic and algorithmic parsing to reduce the burden of work at hand. Second, the divergence of L1 and L2 processing behaviours during comprehension may originate from various factors around L2 learners (e.g., L2 input, L1–L2 interface, task types), which are assumed to anchor the noisier representations of L2 knowledge.

Keywords: Good-enough processing, L2 textbook, L1 property, Task type

Introduction

Second language (L2) knowledge, which is complex and multifaceted (cf. Ellis, 2006), is described as noisier representations in a learners' cognitive space than those involving how first language (L1) knowledge is constructed (Futrell & Gibson, 2017; Tachihara & Goldberg, 2020). This is ascribed to the between-language competition (Frenck-Mestre et al., 2019; Park & Kim, 2021) and increased cognitive load in executing language behaviour (Jacob & Felser, 2016; Pozzan & Trueswell, 2016), resulting in L2ers' reduced capacity to deploy the target knowledge (e.g., Hopp, 2018; Robenalt & Goldberg, 2016) and their learning trajectories distinctive from L1 acquisition (e.g., Jiang et al., 2011; Slabakova, 2014; Shin & Park, 2021). Various proposals have been made in explaining this aspect of L2 knowledge: learners' difficulty in accessing fully specified syntactic structures (e.g., Clahsen & Felser, 2006; but see Omaki & Schulz, 2011), their reduced ability to integrate syntactic representations and information from other cognitive domains (e.g., Sorace, 2011), memory operations during information retrieval for the task at hand (e.g., Hopp, 2014; McDonald, 2006), the extent to which properties of L2 overlap in (and compete with) those of L1 given the constant L1 influence during L2 activities (e.g., Hartsuiker et al., 2004; MacWhinney, 2008).

This study investigates how L2 comprehension proceeds through the lens of the 'good-enough' processing account. Research has shown that, to arrive at a complete interpretation, the linguistic processor (i) analyses linguistic input immediately and incrementally and (ii) selectively involves reanalysis/revision of provisional interpretation only if the previous interpretation goes against the current input (e.g., Altmann & Kamide, 1999; Frazier & Rayner, 1982; Friederici et al., 2001). In this regard, one influential account on how the processor copes with the incoming input maintains that the processor inherently prefers less effortful analysis available at the earliest opportunity over costly computation in real-time processing (e.g., Christianson, 2016; Ferreira, 2003). When the processor interacts with various cues in processing, not all cues are equally influential: the degree and manner that these cues affect the 'good-enough' processing is asymmetric, which is modulated by various factors such as language-specific properties, task demands, cognitive load, and individual differences (e.g., Ferreira & Patson, 2007; Lim & Christianson, 2013; Swets et al., 2008; Tan & Foltz, 2020). Most L1 studies concern only few languages such as English (e.g., Dwivedi, 2013; Kharkwal & Stromswold, 2014); furthermore, studies extending this topic to various L2-learning contexts are scarce. This research bias calls for the need to check whether the previous findings based these languages are generalisable to other languages and language-learning contexts.

We ask how L2 learners with contrastive L1 backgrounds achieve 'good-enough' comprehension. We employ Korean as an L2 and two languages (Czech; English) as learners' L1s, which are typologically distinctive from both Korean and each other. Czech and English are SVO languages but differ in their morpho−syntactic properties. Czech, a synthetic and highly inflectional language, is characterised as an active agreement system through word inflection indicative of grammatical case, gender, and number. This allows flexible word order while keeping propositional meaning intact. English, on the other hand, is an analytic language with little inflection. English has a less active agreement system and rigid word order, and a word's position in a sentence provides crucial information about its grammatical status.

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

Heuristic versus algorithmic processing

According to the 'good-enough' processing account, sentence comprehension proceeds with two parsing routes: algorithmic parsing and heuristic parsing (Christianson, 2016; Ferreira, 2003; Traxler, 2014). Algorithmic parsing yields an analysis based on structural cues extracted from the current input. In this way, the processor computes precise syntactic representations in a strict, bottom-up manner, thereby requiring deep and time-consuming processing. Compared to this costly computation, heuristic parsing facilitates rapid and less effortful (but sometimes incorrect) interpretation based on a comprehender's prior beliefs and knowledge about the incoming input. This parsing route operates mainly on heuristics, which comprise computationally less costly and more accessible options drawn from memory. These "fast and frugal" (Gigerenzer et al., 1999:14) options, which depend upon information other than pure syntactic representations such as ready-made templates (e.g., Townsend & Bever, 2001) and usage frequency (Ambridge et al., 2015; Goldberg, 2019), often return an interim analysis.

Existing literature favouring this account has shown that the processor does not apply the two routes to sentence processing simultaneously or in a balanced way (e.g., Dwivedi, 2013; Ferreira, 2003; Ferreira & Patson, 2007; Kharkwal & Stromswold, 2014). These studies point to the linguistic processor's preference for heuristics over algorithms unless necessary, and thus 'good-enough', as a least-effort strategy for comprehension. One motivation that drives the processor this way occurs when it copes with incoming input at the earliest opportunities. Processing new items invites cognitive challenge (as a form of online disequilibrium), so the processor prefers to restore cognitive equilibrium quickly and to remain in this state for as long as possible (online cognitive equilibrium hypothesis; Karimi & Ferreira, 2016). This leads the processor to prefer heuristic parsing over algorithmic parsing due to an economic advantage of heuristics conserving cognitive effort (although they are provisional and sometimes inaccurate).

Compared to L1 research in 'good-enough' processing, literature on L2 'good-enough' processing is less active. Lim and Christianson (2013), for example, find that L2 processing is strategically 'good-enough'. They measured native English speakers' and L1-Korean L2-English learners' sentence comprehension through self-paced reading and translation, by employing (im)plausible subject/object relative clause sentences. Results showed that the L2 participants were able to use syntactic information to arrive at complete interpretation of a sentence, like the native speaker participants. Notably, the L2 participants' performance was contingent upon proficiency and reading goals, suggesting that L2 comprehension is intertwined with various factors surrounding L2 learners. Tan and Foltz (2020) add to the evidence that task demands affect the extent to which L2 processing becomes 'good-enough'. Native speakers and Chinese-speaking learners of English joined self-paced reading experiments with globally ambiguous and disambiguated (through high- or low-attachment) English relative clause sentences. Task demands were modulated by means of the types of comprehension questions appearing after reading each sentence. Results showed that, like the native speakers, the L2 participants processed the disambiguating region faster when they received superficial questions than when they received questions about relative clause. Moreover, both L1 and L2 participants demonstrated shallow processing, which did not stem from their inability to process the sentences in more detail. The researchers argued that L2 sentence processing is 'good-enough', becoming strategic in that the processor engages in shallow processing when deep processing is not necessarily required for the task.

Suffixal passive in Korean

A suffixal passive (SP) consists of two arguments with the atypical alignment between thematic roles and case-marking (theme–nominative; agent–dative) and passive verbal morphology (Sohn, 1999). A canonical SP follows the theme–agent ordering as in (1a); the verb can be fronted, yielding a verb-initial pattern as in (1b).[1]

(1a) SP: verb-final
totwuk-i kyengchal-hanthey cap-hi-ess-ta.
thief-NOM police-DAT catch-PSV-PST-SE[2]
'The thief was caught by the police.'

(1b) SP: verb-initial
cap-hi-ess-ta totwuk-i kyengchal-hanthey.
catch-PSV-PST-SE thief-NOM police-DAT
'The thief was caught by the police.'

For the verb-final pattern (1a), passive morphology serves as a late-arriving algorithmic cue for identifying voice. The arguments' thematic roles in this pattern are incongruent with the typical composition—nominative-marked agent and accusative-marked theme. Therefore, a reader must revise the initial interpretation by realigning the pairings between thematic roles and case-marking (a theme role to the nominative-marked entity; an agent role to the dative-marked entity), respecting a valency-decreasing process in which a transitive verb loses one argument slot. Notably, these associations, which are exclusive for the passive, are atypical and competes with much stronger and more frequent ones: agent–nominative and recipient–dative. In contrast, the verb-initial pattern (1b) involves a fronted verb (and passive morphology) as an early-arriving algorithmic cue. Therefore, despite having the same anomaly concerning the thematic roles of the case-marked arguments found in the verb-final pattern, this morphology would guide the following interpretation (e.g., Pozzan & Trueswell, 2015), allowing immediate disambiguation of the pairings of the thematic roles and case-marking facts. This would generate a

[1] We controlled for the thematic role ordering (agent-before-theme) and verb location (verb-final versus verb-initial) to compare these patterns with the fewest structural differences.

[2] Abbreviations: DAT = dative marker; N = noun; NOM = nominative case marker; PST = past tense marker; PSV = passive suffix; SE = sentence ender; V = verb.

processing benefit which is not present in the late-arriving cue in achieving a complete interpretation of a sentence.

In sum, the two patterns differ in their respective heuristics (verb-final: frequent word order vs. verb-initial: infrequent word order) and algorithms (verb-final: late-arriving cue → revising the initial interpretation vs. verb-initial: early-arriving cue → guiding the following interpretation; both involving the valency-decreasing process), given the same cue for this computation (passive morphology) and the unusual mapping between thematic roles and case-marking.

Composition of L2 textbook input

It is well-known that L2 textbooks provide L2 learners with an essential/primary source of L2 input (e.g., Alsaif & Milton, 2012; Römer, 2004). As an explanatory purpose, we analysed two textbook types, each of which is used widely in the Czech Republic (Textbook X) and the United States (Textbook Y) for tertiary-level instruction of Korean. Each type consists of multiple proficiency levels, and we selected the first four volumes. This selection process was informed by the actual range of levels that tertiary institutions typically cover in undergraduate language courses. We electronically compiled all the sentences in the textbooks, organising them by textbook type, and then manually extracted the sentences relevant to the investigation. The inclusion/exclusion criteria were the following: each sentence was included if a predicate was overtly realised in a sentence; any sentence was excluded either with no overt predicate present or with a predicate presented in an infinitive form for practice. We finally examined all instances of disagreement and resolved them.

As Table 1 illustrates, the use of SP was scarce, indicating that learners formally instructed with these textbooks in class are not sufficiently exposed to the L2 input regarding SP. The ratios of SP use were asymmetric in such a way that Textbook Y showed a proportionally higher SP use than Textbook X. In contrast, around one-third of the sentences in these textbooks began with a case-marked subject. This number is large considering the major setting of textbook contents—colloquial conversations where sentential components (e.g., particle, case-marked argument) are frequently omitted (Sohn, 1999) or utterances often begin with a vocative case (e.g., *Mia-ya, annyeng* 'Mia, hello'). Moreover, all sentences in these textbooks occurred with the late-arriving predicate. These characteristics may allow learners to get an initial fix on the typical composition of Korean sentences.

Table 1. L2 textbook input composition

Textbook	Total	Instance: # (%)		
		SP	Case-marked subject-first	Predicate-final
X	4,272	9 (0.21)	1,622 (37.97)	4,272 (100.00)
Y	8,759	77 (0.88)	3,077 (35.13)	8,759 (100.00)

Experiment

The scarcity of L2 textbook input regarding SP allows us to explore how L2 comprehension proceeds at the interface of the language-general processing architecture (heuristic-before-algorithm), language-specific properties (word order, case-marking, verbal morphology), and particular tasks at hand. In what following, we measure L2 learners' comprehension and processing through acceptability judgement (AJ) and self-paced reading (SPR) tasks.

In AJ, a comprehender engages in both partial and holistic considerations of a sentence to arrive at a complete interpretation of the sentence and then to reach a verdict on its acceptability. For this task, L2 learners would attend primarily to the typicality of sentence composition in the textbooks (subject-first and predicate-final), which is also consistent with the basic word order in Korean and thus readily available, rather than knowledge specific to the target constructions (verbal morphology and the interpretive procedures driven by that morphology), which is scant in the L2 textbook input and thus laborious to utilise. This should cause the L2 learners' dispreference for a verb-initial pattern relative to a verb-final counterpart, independently of construction type. Considering the robust influence of L1 properties on L2 acquisition (e.g., Frenck-Mestre et al., 2019; Park & Kim, 2021), we also expect L1-Czech L2-Korean learners to be more adept at scrambling than L1-English L2-Korean learners, demonstrating their leniency with the verb-initial pattern than L1-English L2-Korean learners.

In SPR, a comprehender engages in a moment-by-moment and cumulative interpretation of the incoming items. Based on this characteristic, we predict three aspects pertaining to L2 comprehension in this task. First, the learners should spend more time reading sentences than native speakers, as proven by previous studies that show L2 learners' general difficulty in online processing (e.g., Clahsen & Felser, 2006; Hopp, 2014; McDonald, 2006; Pozzan & Trueswell, 2016). Second, the learners would attend more to item-specific requirements within one region currently being handled than to a broad structure-building process by actively utilising cross-region and upcoming information. This is achieved by way of two routes: the learners' capacity for algorithmic parsing driven by morpho–syntactic cues (e.g., Lim & Christianson, 2013); reduced space for generating expectations of the following representations based on the here-and-now information because of the L2 mind being busy integrating information at hand (cf. Grüter et al., 2017). Therefore, these two routes would lead them to cope with the region(s) with verbal morphology, increasing reading times (RTs) at the verb-related regions compared to the same regions of comparison for each construction. Their processing behaviour should contrast to how native speakers, who are fully aware of both heuristics and algorithms involving each construction, would manage a verb-initial pattern relative to a verb-final counterpart, demonstrating the heuristic-before-algorithm processing architecture. Third, like AJ, L1 properties will affect the learners' processing, leading L1-Czech L2-Korean learners to better handle the verb-initial patterns in each construction type with less processing cost than L1-English L2-Korean learners.

Methods

Participants We recruited L1-Czech (CZH) and L1-English (USA) learners of Korean, all of whom were non-heritage

speakers of Korean and currently residing in their respective home countries. Proficiency in Korean was measured through Korean C-test (Lee-Ellis, 2009), for which the main assessment involved participants' comprehension of Korean sentences of varying length and complexity. The two tasks in this study require a certain level of proficiency, so we selected the appropriate participants with the test scores—those with a threshold value of 63 out of 188. In the end, we obtained 28 CZH (M_{age} = 24.1, SD = 2.8; duration of learning Korean: M = 3.9 years, SD = 1.8) and 24 USA (M_{age} = 23.3, SD = 4.2; duration of learning Korean: M = 4.4 years, SD = 2.0) participants. There was no statistical by-group difference in the scores, as proved by the Wilcoxon rank-sum test (W = 399, p = .251). We also recruited 40 native speakers of Korean (NSK) for the control group (M_{age} = 23.6, SD = 4.1).

Stimuli We created 16 test sentences for SPR by canonicity (verb-final; verb-initial). Each test sentence included an invariant carrier phrase, followed by the critical structure and an adverbial phrase consisting of two words (Table 2). We focused on R2, R3, and R4 as the main regions of interest and R5 to accommodate any spill-over effects arising from a task-specific button-press strategy. For subjects in the test sentences, we used simple/common nouns often used in daily life. There was no overlap in verb use across the sentences. The dative marker used in all the sentences was uniformly -*hanthey*, considering that it is used more frequently than -*eykey* in colloquial settings.

Table 2. Scheme of stimuli (SPR)

Pattern	R1	R2	R3	R4	R5	R6
Verb-final	*I heard*	N-NOM	N-DAT	V-PSV	*last night*	
Verb-initial	*that*	V-PSV	N-NOM	N-DAT		

To mask the experiment's intention, the test stimuli were interspersed with 48 fillers in varying structures and complexity. All the test stimuli and fillers were split into four sub-lists and were randomly assigned to participants; we also randomised the sentences' presentation order in each sub-list. Participants also completed AJ to assess their offline knowledge about each condition. The test sentences for this task were generated from the critical structure portion (R2 to R4) from the test stimuli and the fillers used in SPR.

Ten native Korean speakers, who did not participate in the experiment, inspected the test stimuli for their grammaticality using a binary scale (i.e., grammatically correct or incorrect). The result confirmed the grammaticality of the stimuli for the verb-final (mean = 100.00%, SD = 0.00) and the verb-initial (mean = 92.50%, SD = 0.26) conditions. Based on the inspectors' comments, the lower score of the verb-initial stimuli compared to that of the verb-final stimuli was not from the sentences' grammaticality but from the inspectors' dispreference for scrambling without contextual motivation.

Procedure The two tasks were completed sequentially on web-based platforms: *PCIbex* (Zehr & Schwarz, 2018) for SPR; *Qualtrics* for AJ. All sentences were displayed individually (i.e., one by one) on a computer screen. SPR was run visually in a non-cumulative moving-window paradigm (Just et al., 1982). Participants silently read sentences as naturally and quickly as possible, pressing a space bar to work through three practice items before proceeding to the main task. Following the final region in each sentence, a comprehension check-up question appeared prompting participants to answer by clicking on one of the two choices. The question was designed simply to maintain the participants' attention to the task, asking about simple facts regarding the sentence being read (e.g., what the sentence was about, what action was done). This task took around 30 minutes for each participant. For AJ, participants rated the acceptability of each sentence as *quickly* as possible with a 6-point Likert scale (0: unacceptable; 5: acceptable) to respond immediately upon encountering the sentences. Once they clicked on the scale and proceeded to the next sentence, they were not allowed to revise their previous judgements. This task took around 20 minutes for each participant.

Analysis Data from SPR were first trimmed by excluding data from participants who failed to pass the comprehension questions (data loss: 6.32%) and outliers (i.e., extremely long or short RTs) per region through a 3SD cut-off point (data loss: 2.10%). We then log-transformed the trimmed data for data normalisation and further residualised the data to adjust for variability in word length and individuals' reading speed (cf. Baayen & Milin, 2010). Finally, the pre-processed data were submitted to a linear mixed-effect modelling per region, with *Group* and *Canonicity* 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, 2021), including the maximal random-effects structure allowed by the model (Barr et al., 2013). Data from AJ were also trimmed first before the analysis by excluding individual judgement values whose response times were less than 1,000 ms or more than 10,000 ms (data loss: 6.79%). We then Z-transformed the data to satisfy the assumption of normal distribution and proceeded to the same kind of linear mixed-effect modelling, with *Group* and *Canonicity* 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, 2021), including the same kind of maximal random-effects structure as the self-paced reading data (Barr et al., 2013).

Results

AJ (Figure 1): The L2 groups judged the acceptability of all the conditions to be lower than NSK, with the highest around three out of five at most. They also evaluated the verb-initial condition less acceptable than their verb-final counterpart, like NSK. We then conducted pairwise comparisons (NSK−CZH; NSK−USA; CZH−USA) to address group differences. The first two models revealed main effects of *Group* and *Canonicity* (all ps < .001), indicating the L2ers' significantly lower acceptability rates of the verb-initial condition than NSK's and their strong dispreference for the verb-initial condition relative to the verb-final one. The CZH−USA model revealed a main effect of *Canonicity* (p < .001), confirming the uniform gap in rating between the two conditions for both L2 groups.

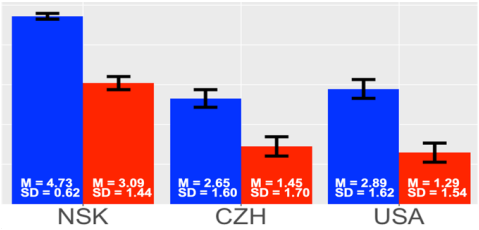

Figure 1. AJ. X-axis: group; Y-axis: acceptability (out of five). Error bars indicate 95% CI.

SPR (Figure 2): For NSK, we found numerically longer RTs in the verb-initial condition than in the verb-final condition except R4 in which passive morphology exerts the assumed revision of the initial interpretation in its typical position. For the L2 groups, we found two general tendencies in the critical (R2 to R4) and spill-over (R5) regions. First, they spent more time at these regions than NSK. Second, whereas NSK showed a constant trend in RTs as reading proceeded, the L2 learner groups demonstrated a sudden drop in RTs at the spill-over region.

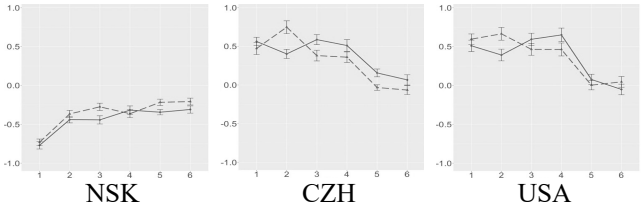

Figure 2. SPR. X-axis: region; Y-axis: RT (residualised). Solid lines indicate verb-final conditions and dotted lines indicate verb-initial conditions. Error bars indicate 95% CI.

We again compared pairs to assess group differences, focusing on the critical and spill-over regions (Table 3). The NSK−CZH model revealed a main effect of *Group* in every region of interest, indicating that CZH globally spent more time than NSK in reading these regions. We found a main effect of *Canonicity* at R2 and an interaction effect at R2, R3, and R5. Post-hoc analyses (α = .025) revealed two points. First, NSK spent significantly more time reading the verb-initial condition than the verb-final condition at R3 (p = .022) and R5 (p = .005). Second, while CZH spent more time reading the verb-initial condition than the verb-final pattern at R2 (p = .011), they took more time reading the verb-final condition than the verb-initial condition at R5 (p = .002). Whereas no statistical significance was found in the RT difference between the verb region (R2) in the verb-initial condition and the verb region (R4) in the verb-final condition for NSK, there was a marginally significant difference in RTs between the two regions for CZH (p = .035).

The NSK−USA model revealed a main effect of *Group* in every region of interest, indicating that USA globally spent more time than NSK in processing these regions. We also found a main effect of *Canonicity* at R2 and an interaction effect at R3 and R5, but none of the post-hoc analyses (α = .025) revealed significance. For USA, there was no statistical significance in RT difference between the verb region (R2) in the verb-initial condition and the verb region (R4) in the verb-final condition.

The CZH−USA model revealed a main effect of *Canonicity* at R2 and R5. Additional analyses (α = .025) revealed significant differences at R2 (p = .011) and R5 (p = .002) only for CZH.

Table 3. By-region statistical model (α = .05)

			β	SE	t	p
NSK−CZH	R2	(Intercept)	−0.030	0.033	−0.890	.389
		Group	0.970	0.054	17.982	< .001***
		Canonicity	0.178	0.065	2.756	.014**
		Grp× Cn	0.273	0.107	2.536	.012**
	R3	(Intercept)	−0.022	0.028	−0.804	.422
		Group	0.846	0.058	14.658	< .001***
		Canonicity	0.018	0.057	0.320	.749
		Grp× Cn	−0.374	0.116	−3.241	.001**
	R4	(Intercept)	−0.038	0.034	−1.112	.282
		Group	0.776	0.060	12.801	< .001***
		Canonicity	−0.088	0.068	−1.284	.217
		Grp× Cn	−0.104	0.121	−0.861	.390
	R5	(Intercept)	−0.142	0.022	−5.949	< .001***
		Group	0.352	0.040	8.708	< .001***
		Canonicity	0.004	0.039	0.109	.913
		Grp× Cn	−0.315	0.079	−3.972	< .001***
NSK−USA	R2	(Intercept)	−0.102	0.034	−3.033	.007**
		Group	0.957	0.058	16.383	< .001***
		Canonicity	0.136	0.053	2.549	.022*
		Grp× Cn	0.207	0.113	1.841	.067
	R3	(Intercept)	−0.060	0.033	−1.794	.098
		Group	0.890	0.062	14.303	< .001***
		Canonicity	0.067	0.066	1.010	.328
		Grp× Cn	−0.299	0.124	−2.403	.017*
	R4	(Intercept)	−0.033	0.037	−0.885	.382
		Group	0.910	0.067	13.626	< .001***
		Canonicity	−0.093	0.061	−1.524	.129
		Grp× Cn	−0.138	0.129	−1.071	.286
	R5	(Intercept)	−0.167	0.027	−6.191	< .001***
		Group	0.334	0.048	6.897	< .001***
		Canonicity	0.059	0.045	1.331	.185
		Grp× Cn	−0.202	0.094	−2.161	.032*
CZH−USA	R2	(Intercept)	0.534	0.048	11.039	< .001***
		Group	−0.042	0.071	−0.595	.553
		Canonicity	0.321	0.097	3.317	.004**
		Grp× Cn	−0.067	0.143	−0.467	.641
	R3	(Intercept)	0.519	0.048	10.759	< .001***
		Group	0.054	0.072	0.745	.457
		Canonicity	−0.166	0.079	−2.096	.053
		Grp× Cn	0.070	0.132	0.528	.598
	R4	(Intercept)	0.490	0.039	12.484	< .001***
		Group	0.122	0.077	1.572	.118
		Canonicity	−0.167	0.078	−2.128	.050
		Grp× Cn	−0.037	0.155	−0.239	.812
	R5	(Intercept)	0.060	0.029	2.090	.044*
		Group	−0.019	0.054	−0.344	.731
		Canonicity	−0.137	0.052	−2.651	.009*
		Grp× Cn	0.112	0.104	1.077	.283

Discussions and Conclusion

The L1 group's performance is notable in two ways. First, they dispreferred the scrambled condition (infrequent and less plausible/felicitous) over the canonical counterpart (frequent and context-neutral). Second, they spent more time reading the verb-initial than the verb-final condition at R5, the spill-over region in which the unusual distributional (induced by word order) and local (induced by the pairings of thematic roles and case-marking) cues are integrated into a complete interpretation of a sentence. These findings suggest that the NSK's comprehension/processing was determined more by the canonicity of word order and the typicality of mapping between thematic roles and case-marking than by the positioning of verbal morphology in a sentence. This points to the selective manifestation of the early-arriving-cue effect (an instance of deep and costly computation) during comprehension, which is contingent upon a speaker's language usage experience, the nature of structural properties dedicated to constructions, and particular language activities required by task types. This aligns with the 'good-enough' processing account that argues the processor's inherent predisposition to favour the heuristic parsing over the algorithmic parsing (e.g., Christianson, 2016; Ferreira, 2003), supporting the idea that the processor operates in a way that maximises cognitive equilibrium in real-time processing (Karimi & Ferreira, 2016).

Patterns of L2 comprehension differed from those of L1 comprehension in this study. For AJ, the L2 groups generally rated two construction types less acceptable than NSK, which resembles the scarcity of SP in the L2 textbooks. However, they dispreferred the verb-initial conditions compared to the verb-final conditions as NSK did, indicating that L2 learners were able to manifest broad word-order sensitivity such that they uniformly rejected the scrambled conditions, regardless of construction type. Recall that many sentences in the textbooks began with a case-marked subject and all sentences ended with a predicate. This may have prompted the learners to form a default clausal strategy with an early-appearing case-marked subject and a late-arriving predicate (which also conforms to the canonical word order in Korean). The rarity of L2 input regarding SP may thus have allowed more space for the word-order heuristic, which is frequent and reliable, to guide the L2 learners' acceptability judgement of the sentences. This supports the role of heuristics, which are essentially frequency-driven, in L2 comprehension.

For SPR, the L2 groups generally exhibited increased RTs when a region involved verbal morphology (verb-initial > verb-final at R2; verb-initial < verb-final at R5 except USA). Processing a verb is more demanding than processing a noun due to the complex nature of a verb with rich morpho–syntactic information used for the structure-building process (e.g., Grimshaw, 1990; Pinker, 1989), and this may have resulted in the RT gap at R2. However, this gap was substantial only for CZH, and we could not find the considerable RT gap at R4; instead, we found that gap in a substantial manner at R5 only for CZH. This requires more explanations on the learners' RT patterns than those based merely on the verb–noun asymmetry.

The findings from SPR bear two possible interpretations. First, the L2 learners may have been forced to conduct the algorithm-based analysis driven by the verbal-morphology cue, rather than the heuristic-based analysis, to go about moment-by-moment and cumulative interpretation. As the L2 processor confronts information that is not strongly supported by optimal and readily available processing strategy—heuristic, it exerts costly and detailed computation involving a verb (and its morphology) to restore cognitive equilibrium. This may have increased RTs at R2 in the verb-initial condition relative to the verb-final condition. The same computational burden arises at R4 (numerically but insignificantly) and at R5 (significantly but selectively) in the verb-final condition. These regions in the verb-final condition are disadvantageous for the L2 processor since it must integrate all the previous and current information to achieve a full clausal interpretation, together with conducting necessary treatment required by passive morphology. This may have invited longer RTs at these regions in the verb-final condition relative to the verb-initial conditions.

The other possibility concerns the asymmetric RT difference between the two conditions by group. At R2 and R5, the RT gap was statistically significant only for CZH. We ascribe it to the interplay between the nature of L2 textbook input and the properties of learners' L1. SP was scarce in both textbook types, but it appeared more in Textbook Y (learnt by USA) than in Textbook X (learnt by CZH). The enhanced presentation of SP in Textbook Y (albeit infrequent) may have assisted USA, whose L1 manifests fixed word order and little use of inflectional morphology, in managing the verb-final condition, relaxing this group's processing load at R5. However, the USA's capacity to cope with the verb-initial condition appears to be still limited, necessitating more processing resources, as shown by their RT patterns at R2.

Together, our results from AJ and SPR suggest that L2 processing does not qualitatively differ from L1 processing. Considering that both heuristic and algorithmic parsing routes are activated in parallel, the L2 processor tends to prioritise heuristics, computationally simpler and cheaper options, than algorithms, computationally more complex and costly options (Lim & Christianson, 2013; Omaki & Schulz, 2011; Tan & Foltz, 2020), like how L1 processor copes with the incoming and upcoming input (e.g., Ferreira, 2003; Kharkwal & Stromswold, 2014). The computational benefit regarding cognitive effort arising from heuristics often makes it easier for the processor to maintain cognitive equilibrium in online sentence processing. Nevertheless, whenever necessary, the L2 processor is willing to engage in algorithmic parsing to enter and restore cognitive equilibrium as immediately and quickly as possible (Karimi & Ferreira, 2016). In addition, various factors such as distributional properties of L2 input, (morpho–syntactic) properties of learners' L1, and particular language activities required by task types jointly modulate the degree to which the L2 processor adjusts its mode to heuristic or algorithmic parsing. We believe the implications of this study's findings advance the current understanding of how these factors contribute to the noisier representations of L2 knowledge.

References

Alsaif, A., & Milton, J. (2012). Vocabulary input from school textbooks as a potential contributor to the small uptake gained by English as a foreign language learners in Saudi Arabia. *Language Learning Journal*, *40*(1), 21–33.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264.

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language, 42*(2), 239–273.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

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.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3-42.

Dwivedi, V. D. (2013). Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PloS One, 8*(11), e81461.

Ellis, N. C. (2006). Cognitive perspectives on SLA: The associative-cognitive CREED. *AILA Review, 19*(1), 100–121.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–203.

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass, 1*(1-2), 71–83.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*(2), 178–210.

Frenck-Mestre, C., Kim, S. K., Choo, H., Ghio, A., Herschensohn, J., & Koh, S. (2019). Look and listen! The online processing of Korean case by native and non-native speakers. *Language, Cognition and Neuroscience, 34*(3), 385–404.

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11*(2), 305–323.

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

Gigerenzer, G., Todd, P. M., & ABC Research Group (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Goldberg, A. E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton University Press.

Grimshaw, J. B. (1990). Argument Structure (Linguistic inquiry monographs 18). Cambridge, MA: MIT Press.

Grüter., T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism, 7*, 199–229.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15*, 409–414.

Hopp, H. (2014). Working memory effects in the L2 processing of ambiguous relative clauses. *Language Acquisition*, *21*(3), 250–278.

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, *17*, 5–27.

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning*, *61*(3), 940–967.

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.

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.

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

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

Lim, J. H., & Christianson, K. (2013). Second language sentence processing in reading for comprehension and translation. *Bilingualism: Language and Cognition*, *16*(3), 518–537.

MacWhinney, B. (2008). A unified model. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 341–371). New York, NY: Routledge.

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.

Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition*, *33*(4), 563–588.

Park, S. H., & Kim, H. (2021). Cross-linguistic influence in the second language processing of Korean morphological and syntactic causative constructions. *Linguistic Approaches to Bilingualism.*

Pinker, S. (1989). *Learnability and cognition. The acquisition of argument structure*. Cambridge, MA: MIT Press.

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

Pozzan, L., & Trueswell, J. C. (2016). Second language processing and revision of garden-path sentences: a visual word study. *Bilingualism: Language and Cognition*, *19*(3), 636–643.

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robenalt, C., & Goldberg, A. E. (2016). Nonnative speakers do not take competing alternative expressions into account the way native speakers do. *Language Learning*, *66*(1), 60–93.

Römer, U. (2004). Comparing real and ideal language learner input: The use of an EFL textbook corpus in corpus linguistics and language teaching. In G. Aston, S. Bernardini, & D. Steward (Eds.), *Corpora and language learners* (pp. 151–168). Amsterdam: John Benjamins.

Shin, G-H., & Park, S. (2021). Isomorphism and second language acquisition: Comprehension of Korean suffixal passive for adult Mandarin-speaking learners of Korean. *Applied Linguistics Review*.

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

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition, 36*(1), 201–216.

Tachihara, K., & Goldberg, A. E. (2020). Reduced competition effects and noisier representations in a second language. *Language Learning, 70*(1), 219–265.

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*, 575847.

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