Studies in Second Language Acquisition, **42** (2020), 1121–1136
doi:10.1017/S0272263120000170

Research Report

SECOND LANGUAGE USERS EXHIBIT SHALLOW MORPHOLOGICAL PROCESSING

Yoonsang Song *

University of Hong Kong

Youngah Do *

University of Hong Kong

Arthur L. Thompson

University of Hong Kong

Eileen R. Waegemaekers

University of Hong Kong

Jongbong Lee

Nagoya University of Commerce & Business

Abstract

The present study tests the Shallow Structure Hypothesis (SSH), which claims that compared to L1 processing, L2 language processing generally underuses grammatical information, prioritizing nongrammatical information. Specifically, this cross-modal priming study tests SSH at the level of morphology, investigating whether late advanced L2 learners construct hierarchically structured representations for trimorphemic derived words during real-time processing as native speakers do. Our results support SSH. In lexical decision on English trimorphemic words (e.g., *unkindness* or [[un-[kind]]-ness]), L1 recognition of the targets was facilitated by their bimorphemic morphological-structural constituent primes (e.g., *unkind*), but not by their bimorphemic nonconstituent primes (e.g., *kindness*), which were only semantically and formally related to the target. In contrast, L2 recognition was equally facilitated by both constituent and nonconstituent primes.

The experiment in this article earned an Open Data badge for transparent practices. The materials are available at https://osf.io/sa37y/

* Correspondence concerning this article should be addressed to Yoonsang Song or Youngah Do, Department of Linguistics, University of Hong Kong, Rm 930, 9/F, Run Run Shaw Tower, Centennial Campus, Pokfulam Road, Hong Kong. E-mail: yoonsang@hku.hk (Y.S.) or youngah@hku.hk (Y.D.)

These results suggest that unlike L1 processing, L2 processing of multimorphemic words is not mainly guided by detailed morphological structure, overrelying on nonstructural information.

INTRODUCTION

How do late L2 learners recognize morphologically complex words in real time? An increasing number of studies are addressing this issue, albeit not without conflicting results. Some studies demonstrate that L2 processing of morphologically complex words is mainly guided by morphological structure like L1 processing (e.g., Coughlin & Tremblay, 2015; Diependaele et al., 2011; Freynik et al., 2017). Other studies, however, provide evidence that L2 lexical processing is distinctively less sensitive to morphological structure and more heavily affected by other types of information than L1 processing (e.g., Clahsen & Neubauer, 2010; Heyer & Clahsen, 2015; Jacob et al., 2017).

Recently, such differences between L1 and L2 morphological processing have been accounted for by Clahsen and Felser's (2006, 2017) Shallow Structure Hypothesis (henceforth SSH) (e.g., Clahsen & Neubauer, 2010; Clahsen et al., 2013; Clahsen et al., 2015; Farhy et al., 2018). The SSH assumes that L1 and L2 speakers are equipped with the same processing architecture and mental processing mechanisms, and that two distinct processing routes are available in both L1 and L2: one involving construction of a detailed grammatical representation for the input and the other involving construction of a grammatically shallow representation that lacks full grammatical detail. However, SSH maintains that the shallow parsing route is in general more likely to dominate in L2 than L1, though this tendency can be affected by other factors such as L2 proficiency or the nature of the given linguistic task. In the shallow parsing route, processing of input is mainly guided by nongrammatical information, such as semantic, pragmatic, surface-form, or relevant nonlinguistic information. Hence, SSH attributes the differences between L1 and L2 processing of morphologically complex words to the underuse of information on morphological structure and overreliance on nonstructural information in L2 processing in general.

This study aims to contribute to the investigation on shallow morphological processing in online L2 word recognition with a unique trimorphemic experiment design. Previous studies on L2 processing of morphologically complex words have limited their scope to bimorphemic words (mostly suffixed ones), addressing the research question of whether the L2 processing of bimorphemic words involves morphological decomposition (e.g., *kindness* → *kind* + *-ness*). We broaden the scope of this line of research by using trimorphemic words as our stimuli set. With the assumption that these words are represented in hierarchical structure, as illustrated in Figure 1 (e.g., Lieber, 1980; Selkirk, 1982), examining the processing of such multimorphemic words has an outstanding merit: It allows us to test whether L2 speakers construct representations with a multilevel-nested structure—rather than linear-sequence structures—during real-time lexical processing. Given that structural hierarchy is one of the most fundamental aspects of the underlying representation of language (e.g., Chomsky, 1995; Jacob et al., 2017; Nelson et al., 2017; Song et al., 2019), investigating whether L2 lexical processing involves real-time construction of hierarchical structure appears essential to define the nature of the processing and mental representation of L2s.

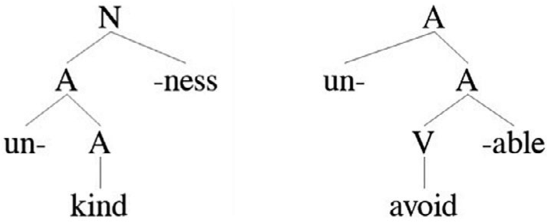

FIGURE 1. Hierarchical structure of the English derived words *unkindness* (left-branching) and *unavoidable* (right-branching).

Note here that bimorphemic words are not ideal for this research question because they have a relatively flat structure. When we assume that these words are also structured in hierarchy with binary branching, the hierarchy only consists of two levels (e.g., [[kind]-ness]). Such hierarchical structure is not clearly distinguished from linear-sequence structure with two elements (e.g., [kind]-[ness]), as both allow for morphological decomposition at the same boundary. In contrast, the hierarchical structure of trimorphemic words (e.g., [[un]-[kind]]-ness]) has three levels of hierarchy and is clearly distinguishable from its counterpart linear-sequence structure (e.g., [un]-[kind]-[ness]). Most importantly, the hierarchical structure specifies which two of the three morphemes combine together first, constituting a morphological-structural constituent (henceforth structural constituent) at the second level of the hierarchy, whereas the linear-sequence structure does not encode such combinatorial process. For instance, the hierarchical structure of the trimorphemic word *unkindness* specifies that to derive the word, the adjective *kind* and the prefix *un-* first merge into an adjective *unkind*, and then the suffix *-ness* attaches to the adjective (see Figure 1). However, in the linear-sequence representation of *unkindness*, it is not encoded whether the root *kind* first combines with the prefix *un-* or the suffix *-ness*, and thus, no such thing as bimorphemic structural constituent exists.

Making use of this three-level hierarchy, recent L1 studies found some empirical evidence that language users build hierarchical structure in real time for trimorphemic words. For example, the results of Bertram et al.'s (2011) eye-tracking study showed that the online processing of trimorphemic compounds slowed down when a hyphen was inserted at the minor constituent boundary (e.g., *foot-ballassociation* for [[foot-[ball]]-association]), whereas the processing did not slow down when such insertion was made at the major constituent boundary (*football-association*).[1] More recently, using a cross-modal priming paradigm, Song et al. (2019) demonstrated that recognition of English trimorphemic derived words (e.g., *unkindness*) was facilitated immediately after their bimorphemic substrings were processed if the substrings were structural constituents of the target words (e.g., *unkind*), and not if they were nonconstituents (e.g., *kindness*). Note that morphological constituents are nested within the morphological structure of targets, whereas nonconstituents are not.

Against this background, the present study tests SSH by investigating L2 processing of trimorphemically derived English words. Specifically, we test whether nonnative English speakers show the same morphological-structural priming as was observed in Song et al. (2019) for native English speakers. In the current study, speakers of L2 English (L1 Cantonese) were tested using the same experimental paradigm and stimuli as

Song et al. (2019). According to SSH, nonnative speakers are likely to exhibit distinct processing patterns from native speakers' in recognition of trimorphemic target words (e.g., *unkindness*). Nonnative speakers' processing is unlikely to differ according to whether target words follow either structural constituent primes or nonconstituent primes (e.g., *unkind* and *kindness*, respectively). Both primes are highly related to the trimorphemic targets semantically and formally (i.e., phonetically and orthographically), and are mainly distinguished by their hierarchical morphological structures. Therefore, SSH allows the possibility for nonnative speakers to show strong priming effects from nonconstituent and constituent primes equally because the shallow processing route, which relies more heavily on semantic and surface-level information than on structural information, is more likely to be adopted over the full parsing route in nonnative processing than in native processing. To the best of our knowledge, the current study is the first to explore L2 processing of multimorphemic words constructed in hierarchy with more than two levels.

METHOD

PARTICIPANTS

Thirty-eight speakers of advanced L2 English were recruited from an English medium university in Hong Kong (mean age = 20.50; *SD* = 1.31). All participants reported Cantonese as their native and dominant language and English as their second language. None considered themselves balanced bilinguals, and almost all participants (97%) reported that they use Cantonese for the most part of their daily life outside school. These L2 learners were considered late learners because they were not substantially exposed to English until they entered secondary school at age 12, although their rather sparse exposure to English occurred early on through English sessions in Cantonese-medium nursery or kindergarten (*M* = 3.11; *SD* = 0.98). Their self-rating on overall English proficiency was 5.63 (*SD* = 0.75) on a scale from 1 (bad) to 7 (native/near-native), while that on overall Cantonese proficiency was 6.89 (*SD* = 0.39). The learners' overall English proficiency was also assessed by administering a C-test taken from Schulz (2006). Their average score was 29.39 out of 40 points (*SD* = 2.34). This score was comparable to the average score of 30.95 (*SD* = 3.45) obtained from 22 advanced late L2 learners in Song's (2015) study who took the same C-test, signifying our participants' high proficiency in English. At the same time, the score was distinctively lower than the average score of 36.05 (*SD* = 1.87) obtained from Song's 19 native English speakers, indicating that our participants were not as proficient as native speakers.

MATERIALS AND DESIGN

Table 1 illustrates the experimental design with examples of prime-target sets, showing the two distinct experimental conditions for English trimorphemic words: LEFT-BRANCHING and RIGHT-BRANCHING (see Figure 1 for their structural differences). A complete list of prime-target sets and their frequency information can be found in the Appendix. All targets were structurally (and semantically) unambiguous, excluding ambiguous words such as *unlockable*. They were matched across branching conditions on frequency and

TABLE 1. Experimental conditions, with an example stimulus set

		Prime type		
Branching condition	Target	Constituent	Nonconstituent	Unrelated
Left-branching	*unkindness*	*unkind*	*kindness*	*prolong*
Right-branching	*unavoidable*	*avoidable*	N.A.	*pluralism*

length using the webCELEX database (http://celex.mpi.nl/). Targets had very low frequency (at zero or nearly zero per million words in webCELEX) or were not listed in the lexical database. The frequency of the unlisted targets was coded as zero. The low frequency of targets made it reasonable to assume that they are likely to be recognized through morphological decomposition rather than their whole-word representations. Each left-branching target (e.g., *unkindness*) was paired with three distinct types of primes, namely, CONSTITUENT, NONCONSTITUENT, and UNRELATED primes. The constituent prime was the bimorphemic structural constituent, which consisted of the first two morphemes of the target (i.e., the prefix and root) (e.g., *unkind*). In contrast, the nonconstituent prime was the bimorphemic derived word consisting of the last two morphemes of the target (i.e., the root and suffix) (e.g., *kindness*). Importantly, the nonconstituent prime did not serve as a structural constituent of the target, although it is equally related to the target in terms of semantics and surface form as the constituent prime. The unrelated prime was a bimorphemic derived word, which was not related to the target in any way (e.g., *prolong*). These three types of primes were matched in frequency and length. In any prime set, the frequency of the constituent prime was not higher than that of the nonconstituent to prevent the constituent prime from being at a potential advantage.

The targets in the right branching condition (e.g., *unavoidable*) were paired with only two prime types, namely, CONSTITUENT and UNRELATED primes. This asymmetry between the two branching conditions occurred, because no right-branching trimorphemic word whose prefix and root constitute a legitimate English word (e.g., **unavoid*) was found among structurally unambiguous trimorphemic words. The two prime types in the right-branching condition were matched in frequency and length. The targets in the two branching conditions were also matched in frequency and length. The right-branching condition was created to test how independently of linear position morphological structural priming occurs. Note that in the left-branching condition, the constituent prime overlaps with the target from the beginning, while the nonconstituent prime does so from the middle. The concern was that this linear position difference may act as a confounding factor, unfairly strengthening priming with the constituent in the left-branching condition. As the constituent prime in the right-branching condition overlaps with the target from the middle like the nonconstituent prime in the left-branching condition, adding the right-branching condition allows us to test how independently of linear position morphological-structural priming occurs.

The complete experimental stimulus set consisted of 30 prime-target pairs (18 left-branching sets and 12 right-branching sets). These 30 experimental sets were distributed across six lists using a Latin square design. That is, an experimental target was paired with only one of the different types of primes in each list. Each list also included 70 filler prime-target pairs, and consequently had 100 prime-target pairs in total. Among the 100 targets, half were words, while the other half were nonwords.

PROCEDURE

Participants were tested individually in a quiet laboratory setting using SuperLab version 4.0, a software developed by Cedrus Corporation, for stimuli presentation and data collection. A cross-modal priming paradigm (auditory primes and visual targets) was used because this paradigm has an advantage over a unimodal priming paradigm (e.g., visual-visual): When priming effects are found in a cross-modal paradigm, we can exclude the possibility that the priming occurred through overlap in low-level modality-specific representations (e.g., orthographic priming). Rather, it can be argued that the mediation between primes and targets occurs through their modality-independent central lexical representations in the paradigm (Marslen-Wilson et al., 1994).

In the experiment, participants were asked to perform a lexical decision task on visual targets, pressing the “Word” or “Nonword” button on a response pad as quickly and as accurately as possible. “Word” responses were always made by the dominant hand. Primes were prerecorded by a male native speaker of American English in a soundproof booth. Every trial began with a fixation cross presented in the center of the screen in white color against a black background (500 ms). The fixation cross was followed by an auditory prime, which was presented over headphones. At the acoustic offset of the prime, a visual target was displayed in lowercase 26-point Arial font in the center of the screen. The target remained there until a lexical decision was made or up to a timeout of 2000 ms. The between-trial intervals were 1000 ms. The presentation order of prime-target pairs was randomized for each participant. Prior to presentation of experimental items, participants were given 10 practice trials (five with word targets and five with nonword targets). After each practice trial, participants received feedback on the accuracy and speed of their response. No feedback was provided to responses during experimental trials.

DATA ANALYSIS

No items were excluded. Data from two participants were excluded from analysis due to their high error rate (> 30%). The data from the remaining 36 participants were merged and compared with data from Song et al.’s (2019) 29 native speakers, who underwent the same task. Response time (RT) analysis included only RTs from experimental targets on which a correct response was made, while accuracy analysis included data from all 30 critical targets.

For statistical analysis, the lme4 (Bates et al., 2015) and lmerTest package (Kuznetsova et al., 2016) were used in the R statistical computing environment. The RT and accuracy data were analyzed with linear mixed-effects regression (Baayen et al., 2008) and with logit mixed-effects regression (Jaeger, 2008), respectively. As a dependent variable, RT models included log-transformed RTs, while accuracy models included log odds of making an incorrect lexical decision. First, we built models for global analysis that included Language Group (L1 vs. L2), Prime Type (Constituent vs. Unrelated), and Branching (Left vs. Right) as fixed effect variables, whereas participant and item as crossed random effects. Then, we fitted models for L1-L2 joint analysis for each branching condition because the number of levels of the Prime Type variable varied in the two branching conditions (left-branching: three; right-branching: two). Recall that the nonconstituent level was included only in the left-branching condition. Furthermore, within-language-group analysis was carried out for L2 data to ensure that any effect found in the L1-L2 joint analysis was not mainly driven by native speakers. Models fitted for this L2 analysis only included Prime Type as a fixed effect, as they were separately fitted for each branching condition.

RESULTS

GLOBAL ANALYSIS INCLUDING BOTH BRANCHING CONDITIONS AND BOTH LANGUAGE GROUPS

Figures 2 and 3 summarize RT and accuracy data, respectively, for the L2 learners in the current study, in comparison to the native speakers in Song et al.'s (2019) study. These figures show that overall, L2 learners responded more slowly and made more errors than

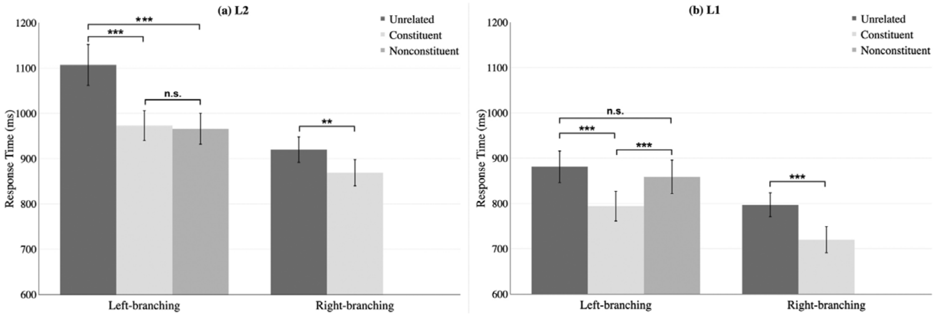

FIGURE 2. The mean response times of the L2 and L1 group in each priming condition of each branching condition. Error bars indicate standard errors. L1 data are from Song et al. (2019). Statistical significance codes: *** = $p < .001$, ** = $p < .01$, * = $p < .05$, n.s. = $p > .05$.

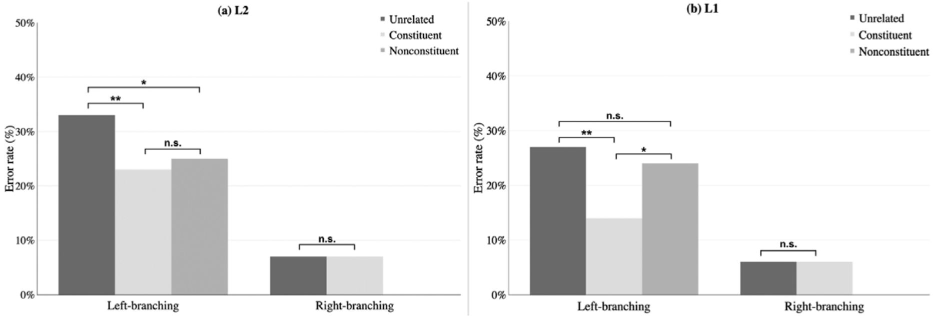

FIGURE 3. The error rates of the L2 and L1 group in each priming condition of each branching condition. L1 data are from Song et al. (2019). Statistical significance codes: *** = $p < .001$, ** = $p < .01$, * = $p < .05$, n.s. = $p > .05$.

TABLE 2. (a) RT and (b) accuracy model for global analysis including both branching conditions and both language groups

(a) RT model

Fixed effects	Beta	Std. Err.	df	*t*-Value	*p*
(Intercept)	6.75	0.03	77.92	229.59	<0.001
Prime type	−0.11	0.01	1219.80	−8.83	<0.001
Language group	0.19	0.05	63.11	4.15	<0.001
Branching	−0.13	0.04	26.63	−3.50	0.002
Prime type × language group	0.02	0.02	1216.66	1.05	0.29
Prime type × branching	0.05	0.02	1219.01	1.92	0.055
Language group × branching	−0.04	0.02	1218.42	−1.79	0.074
Prime type × language group × branching	0.03	0.05	1216.17	−0.71	0.48

(b) Accuracy model

Fixed effects	Beta	Std. Err.	*z*-Value	*p*
(Intercept)	−2.45	0.26	−9.45	<0.001
Prime type	−0.43	0.18	−2.40	0.016
Language group	0.39	0.27	1.44	0.149
Branching	−1.74	0.46	−3.79	<0.001
Prime type × language group	0.28	0.35	0.80	0.422
Prime type × branching	0.82	0.35	2.30	0.021
Language group × branching	−0.31	0.35	−0.87	0.386
Prime type × language group × branching	−0.05	0.71	−0.08	0.939

Note: Fixed-effect factors are sum-coded (prime type: unrelated [−0.5] and related [0.5]; group: L1 [−0.5] and L2 [0.5]; branching: left [−0.5] and right [0.5]).

native speakers, as commonly reported in the L2 processing literature. The linear mixed-effects models fitted for these data confirmed that the group difference in RTs was statistically significant, although that in error rates was not, as the two sections of Table 2 (a and b) respectively show: The main effect of Language Group was significant in the RT model, but not in the accuracy model, indicating that the L1-L2 difference was clearer in terms of RTs.

Figures 2 and 3 also show contrasts between the two branching conditions: Participants in both language groups responded faster and made less errors in the right-branching condition than in the left-branching condition overall, suggesting that right-branching targets were recognized more easily than left-branching targets in both language groups. Table 2 confirms that this difference was statistically significant with the significant main effect of branching.

Importantly, Figure 2 shows that in both language groups, participants' RTs reduced with the constituent primes compared to the unrelated primes in both branching conditions. Figure 3 also shows that across language groups, participants' error rates dropped for the constituent primes compared to the unrelated primes at least in the left-branching condition. The significant main effects of Prime Type (Constituent vs. Unrelated) in Table 2 (a and b) confirm that the priming effect in the constituent condition was statistically significant across language groups and across branching conditions. Note that Figure 3 does not show such priming effects in terms of accuracy in the right-branching condition due to a ceiling effect (i.e., very high accuracy). This difference between the two branching conditions was reflected in the significant interaction between Prime Type and Branching (see Table 2b).

As Table 2 shows, the main effect of Prime Type was not significantly modulated by the Language Group variable in terms of either RTs or accuracy, suggesting that the priming effects in the constituent condition for the two language groups were comparable in magnitude: Either the two-way interaction between Prime Type and Language Group or the three-way interaction among Prime Type, Branching, and Language Group was not statistically significant. That is, the global analysis did not reveal any language group difference in terms of priming in the constituent condition. Lastly, the marginally significant interaction between Prime Type and Branching in the RT model explains the smaller priming effects in the right-branching condition than in the left-branching condition shown in Figure 2 (especially in the L2 group).

Note also that Figures 2 and 3 show that in the left-branching condition, L2 learners' RTs and error rates in the nonconstituent condition were close to those in the constituent condition, whereas native speakers' RTs and error rates in the nonconstituent condition were close to those in the unrelated condition. That is, our descriptive statistics show a clear priming effect in the nonconstituent condition in the L2 group, but not in the L1 group. This apparent group difference could not be statistically tested in the global analysis, as it did not include data from the nonconstituent condition. Therefore, in the next section, we report the results of a separate analysis for each branching condition, in which data from all priming conditions are included.

ANALYSIS FOR EACH BRANCHING CONDITION

RT Analysis

Figure 2 and the global statistical analysis showed that both language groups responded significantly faster in the constituent condition than in the unrelated condition in both branching conditions. The linear mixed-effects model fitted for RT data from each branching condition confirmed again that these priming effects were statistically significant: The main effect of Prime Type (Constituent vs. Unrelated) was significant in both branching conditions as shown in Table 3 (a and b). These priming effects were not modulated by Language Group, as in the global analysis: The interaction between Language Group and Prime Type (Constituent vs. Unrelated) was not significant in either of the branching conditions, as Table 3 shows.

Table 3a also shows that the main effect of Prime Type (Nonconstituent vs. Unrelated) was not significant, suggesting that as a group, the L1 and L2 speakers did not show a significant priming effect in the nonconstituent condition. However, the interaction between Prime Type (Nonconstituent vs. Unrelated) and Language Group was significant, indicating that nonconstituent primes led to a significantly larger RT drop for the L2 group than for the L1 group, compared to unrelated primes. This significant interaction was in line with within-group analyses for L2 (see Table 4a) and L1 data (see Table A1 in Song et al., 2019): The coefficient of Prime Type (Nonconstituent) in the left-branching condition was only significant in the L2 group, indicating that a significant nonconstituent priming only occurred in the L2 group. Furthermore, when we directly compared the mean RTs in the

TABLE 3. Response time model for (a) left- and (b) right-branching condition

(a) Left-branching condition

Fixed effects	Beta	Std. Err.	df	*t*-Value	*p*
(Intercept)	6.81	0.04	43.40	184.78	<0.001
Prime type (constituent vs. unrelated)	−0.12	0.02	797.92	−5.52	<0.001
Prime type (nonconstituent vs. unrelated)	−0.02	0.02	798.13	−1.11	0.267
Language group	0.19	0.05	62.84	3.73	<0.001
Prime type (constituent vs. unrelated) × language group	0.06	0.04	797.06	1.48	0.16
Prime type (nonconstituent vs. unrelated) × language group	−0.11	0.04	797.47	−2.49	0.013

(b) Right-branching condition

Fixed effects	Beta	Std. Err.	df	*t*-Value	*p*
(Intercept)	6.68	0.03	43.67	238.30	<0.001
Prime type (constituent vs. unrelated) × language group	−0.08	0.02	651.67	−5.38	<0.001
	0.17	0.05	62.94	3.75	<0.001
Prime type (constituent vs. unrelated) × language group	0.04	0.03	650.58	1.33	0.185

Note: Fixed-effect factors are sum-coded (prime type in the left-branching condition: unrelated [−0.5, −0.5], constituent [0.5, 0], and nonconstituent [0, 0.5]; group: L1 [−0.5] and L2 [0.5]; prime type in the right-branching condition: unrelated [−0.5] and related [0.5]).

constituent and nonconstituent conditions for the L2 group (by running again the model reported in Table 4a with the reference level of Prime Type switched to Nonconstituent from Unrelated), they did not significantly differ ($\beta = -0.004$, $t(420.19) = -0.15$, $p > 0.10$). In contrast, in the L1 group, the mean RT in the constituent condition was significantly shorter than that in the nonconstituent condition ($\beta = -0.09$, $t(362.50) = -3.44$, $p < 0.001$), as reported in Song et al. (2019). All these results are consistent with Figure 2, where in the L2 group, the mean RT dropped slightly further in the nonconstituent condition than in the constituent condition (compared to the unrelated condition), while in the L1 group, it dropped far less in the nonconstituent condition than in the constituent condition.

Lastly, the within-group analysis confirmed once again that L2 learners' lexical decision times significantly reduced by constituent primes compared to unrelated primes in both branching conditions: Table 4 shows that the coefficients of Prime Type (Constituent) in the two branching conditions were both significant.

Accuracy Analysis

Figure 3 and the global statistical analysis showed that in the left-branching condition, both language groups made less errors in the constituent condition than in the unrelated condition. The L1-L2 joint analysis for each branching condition confirmed again that this priming effect was statistically significant and not modulated by Language Group: Table 5a shows that the main effect of Prime Type (Constituent vs. Unrelated) was significant, but its interaction with Language Group was not. In addition, Figure 3 shows a large facilitation with nonconstituent primes in the left-branching condition for L2

TABLE 4. Response time and accuracy model for (a) left- and (b) right-branching condition for the L2 group

(a) RT model for left-branching condition

Fixed effects	Beta	Std. Err.	df	*t*-Value	*p*
(Intercept)	6.99	0.05	54.55	147.36	<0.001
Prime type (constituent)	−0.13	0.03	421.40	−4.72	<0.001
Prime type (nonconstituent)	−0.13	0.03	422.01	−4.52	<0.001

(b) RT model for right-branching condition

Fixed effects	Beta	Std. Err.	df	*t*-Value	*p*
(Intercept)	6.80	0.04	43.74	183.58	<0.001
Prime type (constituent)	−0.06	0.02	354.40	−2.99	0.003

(c) Accuracy model for left-branching condition

Fixed effects	Beta	Std. Err.	*z*-Value	*p*
(Intercept)	−0.99	0.36	−2.74	0.006
Prime type (constituent)	−0.72	0.25	−2.87	0.004
Prime type (nonconstituent)	−0.53	0.25	−2.14	0.033

(d) Accuracy model for right-branching condition

Fixed effects	Beta	Std. Err.	*z*-Value	*p*
(Intercept)	−3.65	0.71	−5.17	<0.001
Prime type (constituent)	−0.07	0.41	−0.17	0.862

Note: The fixed effect (prime type) is treatment-coded with the reference level set as prime type: unrelated.

learners and a small facilitation for native speakers, but the facilitation was not statistically significant across the language groups: The main effect of Prime Type (Nonconstituent vs. Unrelated) was found statistically not significant, as shown in Table 5a. The interaction between Prime Type (Nonconstituent vs. Unrelated) and Language Group did not reach statistical significance either, unlike in the counterpart RT analysis. As expected by the equally very low error rates of the constituent and unrelated conditions shown in Figure 3, the accuracy analysis for the right-branching condition failed to provide any evidence for a significant priming effect from constituent primes: Table 5b shows that in the right-branching condition, the main effect of Prime Type (Constituent vs. Unrelated) and its interaction with Language Group both failed to reach statistical significance.

Although we did not find any language-group effect on priming, we further analyzed L2 accuracy data separately as planned. The results of this analysis are provided in Table 4 (c and d). First, these tables show that the coefficient of Prime Type (Constituent) was significant in the left-branching condition, but not in the right-branching condition, confirming once again that constituent primes induced significant priming only in the left-branching condition. Crucially, the coefficient of Prime Type (Nonconstituent) turned out to be significant, suggesting that significant nonconstituent priming occurred in the L2 group (see Table 4c). The direct comparison between the constituent and the nonconstituent condition—conducted by rerunning the model reported in Table 4c with the

TABLE 5. Accuracy model for (a) left- and (b) right-branching condition

(a) Left-branching condition

Fixed effects	Beta	Std. Err.	*z*-Value	*p*
(Intercept)	−1.58	0.30	−5.32	<0.001
Prime type (constituent vs. unrelated)	−0.88	0.23	−3.74	<0.001
Prime type (nonconstituent vs. unrelated)	0.07	0.22	−0.32	0.753
Language group	0.40	0.28	1.42	0.156
Prime type (constituent vs. unrelated) × language group	0.58	0.47	1.25	0.212
Prime type (nonconstituent vs. unrelated) × language group	−0.57	0.44	−1.28	0.200

(b) Right-branching condition

Fixed effects	Beta	Std. Err.	*z*-Value	*p*
(Intercept)	−3.26	0.41	−8.00	<0.001
Prime type (constituent vs. unrelated)	0.00	0.31	0.00	0.998
Language group	0.24	0.36	0.68	0.505
Prime type (constituent vs. unrelated) × language group	0.17	0.62	0.27	0.787

Note: Fixed-effect factors are sum-coded (prime type in the left-branching condition: unrelated [−0.5, −0.5], constituent [0.5, 0], and nonconstituent [0, 0.5]; group: L1 [−0.5] and L2 [0.5]; prime type in the right-branching condition: unrelated [−0.5] and related [0.5]).

reference level of Prime Type switched to Nonconstituent from Unrelated—did not reveal a statistically significant difference ($\beta = -0.19, z = -0.75, p > 0.10$). These results contrast with those of the within-group analysis for the L1 group: Significant priming only occurred in the constituent condition (see Table A3 in Song et al., 2019), and the error rate was significantly lower in the constituent than in the nonconstituent condition for native speakers ($\beta = -0.76, z = -2.46, p = 0.014$), as was in the counterpart RT analysis.[2]

DISCUSSION

The present study tested SSH at the morphology level. Specifically, this study investigated whether late advanced L2 learners construct nested morphological structures for multimorphemic words during real-time word recognition, as native speakers do (Song et al., 2019). The SSH predicts that constituent primes and nonconstituent primes yield equally strong priming effects for nonnative speakers because their processing relies more heavily on nonstructural information (such as semantic and surface-form information) than on structural information. Recall that both constituent and nonconstituent primes are highly related to their trimorphemic targets in terms of surface form and semantics, as they both share the root and an affix with the targets, but only constituent primes are nested in the hierarchical morphological structure of targets. Therefore, if L2 processing of multimorphemic words is mainly guided by their morphological structure, constituent primes are expected to yield a larger priming effect than nonconstituent primes; otherwise, they are only expected to produce as strong a priming effect as nonconstituent primes do.

Our results are consistent with SSH, revealing equally strong priming effects from constituent and nonconstituent primes in the L2 group. As Figures 2a and 3a show, in the left-branching condition, L2 learners' mean RTs and error rates in these two conditions dropped to a similar extent compared to those in the unrelated condition, and a series of our statistical analysis confirmed that these priming effects were both statistically significant. Also, direct comparison between the two conditions in terms of RTs and error rates revealed no significant difference. Taken together, all these results suggest that in the L2 group, the priming observed in the constituent condition was mainly driven by the semantic and surface-form similarity between constituent primes and targets, rather than by morphological structure.

These results of nonnative speakers clearly contrast with those of native speakers: Not only did native speakers' mean RT and error rate drop significantly more in the constituent condition than in the nonconstituent condition (see Figures 2b and 3b), only the drop in the constituent condition turned out to be statistically significant. That is, no nonconstituent priming occurred in the L1 group. Song et al. (2019) explained that during the processing of targets, potential processing facilitation led by the semantic and surface-form overlap between targets and their nonconstituent primes might have been cancelled out by processing delay that stemmed from the structural inconsistency between the nonconstituent prime and the target (see, Song et al., 2019, for a more detailed discussion). In light of this interpretation, L1 and L2 processing of morphologically complex words proceed in distinct ways: While L1 processing is so sensitive to structural information that the influence of all nonstructural information is all cancelled out, L2 processing is far more sensitive to nonstructural information than to structural information.

However, this processing difference between L1 and L2 does not necessarily imply that the L2 learners in this study have failed to acquire/develop nativelike representations of the multimorphemic words, where individual morphemes are arranged in a nested hierarchy. Previous studies show that successful acquisition of target grammatical representations (or knowledge) does not automatically lead to nativelike processing patterns, suggesting that L2 speakers have trouble making use of their grammatical representation during online processing (e.g., Felser et al., 2012; Felser & Cunnings, 2012). Therefore, with the results of this study per se, one cannot argue that L2 processing is insensitive to morphological structure simply because hierarchical morphological structure is absent in the L2 representation (or the L2 representation of such structure is seriously flawed in some ways). The SSH acknowledges that both nontargetlike grammatical representations and a nontargetlike processing system can be a potential source of the different grammatical processing patterns between L1 and L2 (Clahsen & Felser, 2017, p. 4). Close examination of the L2 representation of morphological structure is needed in the future to disentangle these two explanations for the distinct patterns of L1 and L2 morphological processing observed in this study.

To summarize, our behavioral data suggest that L2 online recognition of multimorphemic words is mainly guided by nonstructural information rather than their morphological structure. Such L2 processing contrasts with L1 processing, where structural hierarchy plays a major role. Our results are in line with SSH, which claims that L2 learners, even highly proficient ones, tend to underuse grammatical information during real-time language processing compared to native speakers. Future research should further investigate the mental representation of multimorphemic words in L2 to better understand the underlying source(s) of shallow morphological processing.

NOTES

[1] In Dutch and Finnish, multimorphemic words are normally written in concatenated format without a space or a hyphen.

[2] This comparison was not carried out in Song et al. (2019), and thus we conducted it ourselves using the data from the study.

REFERENCES

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 48. https://doi.org/10.18637/jss.v067.i01.

Bertram, R., Kuperman, V., Baayen, R. H., & Hyönä, J. (2011). The hyphen as a segmentation cue in triconstituent compound processing: It's getting better all the time. *Scandinavian Journal of Psychology*, *52*, 530–544. https://doi.org/10.1111/j.1467-9450.2011.00914.x.

Chomsky, N. (1995). *The minimalist program*. MIT Press.

Clahsen, H., Balkhair, L., Schutter, J.-S., & Cunnings, I. (2013). The time course of morphological processing in a second language. *Second Language Research*, *29*, 7–31. https://doi.org/10.1177/0267658312464970.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42. https://doi.org/10.1017/S0142716406060024.

Clahsen, H., & Felser, C. (2017). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*, 1–14. https://doi.org/10.1017/S0272263117000250.

Clahsen, H., Gerth, S., Heyer, V., & Schott, E. (2015). Morphology constrains native and non-native word formation in different ways: Evidence from plurals inside compounds. *The Mental Lexicon*, *10*, 53–87. https://doi.org/10.1075/ml.10.1.03cla.

Clahsen, H., & Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, *120*, 2627–2637. https://doi.org/10.1016/j.lingua.2010.06.007.

Coughlin, C. E., & Tremblay, A. (2015). Morphological decomposition in native and non-native French speakers. *Bilingualism: Language and Cognition*, *18*, 524–542. https://doi.org/10.1017/S1366728914000200.

Diependaele, K., Duñabeitia, J. A., Morris, J., & Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, *64*, 344–358. https://doi.org/10.1016/j.jml.2011.01.003.

Farhy, Y., Veríssimo, J., & Clahsen, H. (2018). Do late bilinguals access pure morphology during word recognition? A masked-priming study on Hebrew as a second language. *Bilingualism: Language and Cognition*, *21*, 1–7. https://doi.org/10.1017/S1366728918000032.

Felser, C., Cunnings, I., Batterham, C., & Clahsen, H. (2012). The timing of island effects in nonnative sentence processing. *Studies in Second Language Acquisition*, *34*, 67–98. https://doi.org/10.1017/S0272263111000507.

Felser, C., & Cuninings, I. A. N. (2012). Processing reflexives in a second language: The timing of structural and discourse-level constraints. *Applied Psycholinguistics*, *33*, 571–603. https://doi.org/10.1017/S0142716411000488.

Freynik, S., Gor, K., & O'Rourke, P. (2017). L2 processing of Arabic derivational morphology. *The Mental Lexicon*, *12*, 21–50. https://doi.org/10.1075/ml.12.1.02fre.

Heyer, V., & Clahsen, H. (2015). Late bilinguals see a scan in scanner AND in scandal: Dissecting formal overlap from morphological priming in the processing of derived words. *Bilingualism: Language and Cognition*, *18*, 543–550. https://doi.org/10.1017/S1366728914000662.

Jacob, G., Heyer, V., & Veríssimo, J. (2017). Aiming at the same target: A masked priming study directly comparing derivation and inflection in the second language. *International Journal of Bilingualism*, *22*, 619–637. https://doi.org/10.1177/1367006916688333.

Jacob, G., Katsika, K., Family, N., & Allen, S. E. M. (2017). The role of constituent order and level of embedding in cross-linguistic structural priming. *Bilingualism: Language and Cognition*, *20*, 269–282. https://doi.org/10.1017/S1366728916000717.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446. https://doi.org/10.1016/j.jml.2007.11.007.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2016). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13.

Lieber, R. (1980). *On the organization of the lexicon*. (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, 3–33. https://doi.org/10.1037/0033-295X.101.1.3.

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S. S., Naccache, L., Hale, J. T., Dehaene, S., & Pallier, C. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, *114*, E3669–E3678. https://doi.org/10.1073/pnas.1701590114.

Schulz, B. (2006). *Wh scope marking in English interlanguage grammars: Transfer and processing effects on the second language acquisition of English complex questions*. (Unpublished doctoral dissertation). University of Hawaii at Mānoa, Mānoa, HI.

Selkirk, E. (1982). *The syntax of words*. MIT Press.

Song, Y. (2015). L2 Processing of Plural Inflection in English. *Language Learning*, *65*, 233–267. https://doi.org/10.1111/lang.12100.

Song, Y., Do, Y., Lee, J., Thompson, A., & Waegemaekers, E. (2019). The reality of hierarchical morphological structure in multimorphemic words. *Cognition*, *183*, 269–276. https://doi.org/10.1016/j.cognition.2018.10.015.

APPENDIX

A complete list of prime-target pairs and their frequency information. The frequency of each word (per Million) in webCELEX (http://celex.mpi.nl/) is provided in parentheses.

Branching condition	Target		Constituent prime		Nonconstituent prime		Unrelated prime	
Left	readjustable	(0)	readjust	(1)	adjustable	(2)	womanhood	(1)
Left	reattachable	(0)	reattach	(0)	attachable	(0)	trickster	(1)
Left	reclosable	(0)	reclose	(0)	closable	(0)	miscount	(0)
Left	recomputable	(0)	recompute	(0)	computable	(0)	sugarless	(0)
Left	rehydratable	(0)	rehydrate	(0)	hydratable	(0)	authorship	(0)
Left	reobtainable	(0)	reobtain	(0)	obtainable	(3)	disinfect	(1)
Left	resealable	(0)	reseal	(0)	sealable	(0)	piggish	(0)
Left	resellable	(0)	resell	(0)	sellable	(0)	subplot	(0)
Left	reusable	(0)	reuse	(1)	usable	(3)	decode	(2)
Left	unawareness	(0)	unaware	(16)	awareness	(24)	pavement	(21)
Left	unclearness	(0)	unclear	(0)	clearness	(0)	outsmart	(0)
Left	uncleverness	(0)	unclever	(0)	cleverness	(2)	breakage	(1)
Left	unhappiness	(6)	unhappy	(28)	happiness	(28)	midnight	(24)
Left	unholiness	(0)	unholy	(1)	holiness	(3)	wastage	(2)
Left	unkindness	(0)	unkind	(4)	kindness	(12)	prolong	(5)
Left	unripeness	(0)	unripe	(0)	ripeness	(1)	erasure	(0)
Left	unsharpness	(0)	unsharp	(0)	sharpness	(2)	postdate	(0)
Left	unwariness	(0)	unwary	(0)	wariness	(1)	draftee	(0)
Right	unavoidable	(4)	avoidable	(1)	NA		pluralism	(1)
Right	unbreakable	(0)	breakable	(0)	NA		misadvise	(0)
Right	undeniable	(2)	deniable	(0)	NA		humidify	(0)
Right	unmovable	(0)	movable	(2)	NA		preview	(2)

Branching condition	Target		Constituent prime		Nonconstituent prime	Unrelated prime	
Right	unreadable	(1)	readable	(1)	NA	ticklish	(1)
Right	untouchable	(1)	touchable	(0)	NA	semifinal	(0)
Right	uneventful	(0)	eventful	(1)	NA	divorcee	(1)
Right	ungraceful	(0)	graceful	(7)	NA	needless	(6)
Right	unresentful	(0)	resentful	(6)	NA	publicize	(7)
Right	unskillful	(0)	skillful	(7)	NA	hardship	(8)
Right	unthankful	(0)	thankful	(4)	NA	interact	(5)
Right	untruthful	(0)	truthful	(4)	NA	virtuous	(5)