Journal of Teaching Language Skills
(JTLS)
39(1), Spring 2020, pp. 153-199
DOI: 10.22099/jtls.2020.37229.2828

Online ISSN
2717-1604

Research Paper

Frequency Effects of Regular Past Tense Forms in English on Native Speakers' and Second Language Learners' Accuracy Rate and Reaction Time

Ebrahim Safaie *

Department of English Language, Islamic Azad University, Lamerd Branch

Abstract

There is substantial debate over the mental representation of regular past tense forms in both first language (L1) and second language (L2) processing. Specifically, the controversy revolves around the nature of morphologically complex forms such as the past tense *–ed* in English and how morphological structures of such forms are represented in the mental lexicon. This study focuses on the results of a speeded acceptability judgment task testing English regular past tense forms of high- and low frequencies. In this task, participants judged the acceptability of sentences as quickly and accurately as possible. Thirty-two intermediate-to-advanced L1 Persian and L1 Arabic speakers (L2ers) and twenty-two Native speakers (NSs) of English made acceptability judgments for regular past tense forms of high- and low-frequency verbs in sentential contexts. Considering participants' reaction times (RTs) and accuracy rates as the dependent variables, the main results are as follows. Despite NSs' faster RTs and higher accuracy scores, both groups demonstrated the same pattern of accuracy rates and RTs. Specifically, for accuracy data, regular verbs yielded a reverse frequency effect or anti-frequency effect in both groups (i.e., lower accuracy rates for high- than low-frequency regular forms in NSs as well as L2ers). For RT data, while the NSs exhibited a marginally anti-frequency effect, the L2ers displayed a nonsignificant trend for the anti-frequency effect. These results support the dual-mechanism models suggesting that the mental mechanisms and representations of inflectional morphology are the same in NSs and intermediate-to-advanced L2ers.

Keywords: Dual-mechanism models, Frequency effect, Mental representation, Regular past tense, Second language

Received: 06/07/2020 Accepted: 07/09/2020

* Assistant Professor, Email: esafai@iaulamerd.ac.ir

The mental representation of regular past tense forms has been the subject of substantial debate in both first language (L1) and second language (L2) processing. (Clahsen & Neubauer, 2010; Pinker, 1999; Silva & Clahsen, 2008; Ullman, 2001; Ullman, 2005). Specifically, much of the debate revolves around the nature of morphologically complex forms such as the past tense –*ed* in English and whether Second Language Learners (L2ers) use the same mechanisms as Native Speakers (NSs) in the processing of such complex forms.

Generally speaking, two broad positions have emerged from research on morphological processing. The first position holds that, like NSs, high proficiency L2ers are sensitive to the morphological structure of complex forms, thereby decomposing them during processing. That is, high proficiency L2ers use similar mechanisms to NSs during the processing of morphologically complex forms (Basnight-Brown, Chen, Hua, Kostić, & Feldmann, 2007; Beck, 1997; Birdsong & Flege, 2001; Feldman, Kostic, & Basnight-Brown, 2010; Hahne, Mueller, & Clahsen, 2006; Lalleman, van Santen, & van Heuven, 1997; Pliatsikas & Marinis, 2013a; Voga, Anastassiadis-Symeonidis, & Giraudo, 2014). Still, very recently, it has been found that not only high proficiency but also low proficiency L2ers pattern alike NSs in their sensitivity to morphological structure (e.g., Coughlin, Fiorentino, Royle, & Steinhauer, 2019).

The second position is that even high proficiency L2ers use different mechanisms from NSs in that they rely on whole-word storage more than decomposition during the processing of morphologically complex forms (Babcock, Stowe, Maloof, Brovetto, & Ullman, 2012; Bowden, Gelfand, Sanz, & Ullman, 2010; Clahsen, Balkhair, Schutter, & Cunnings, 2013; Jacob, Hever, & Veríssimo, 2018; Neubauer & Clahsen, 2009; Pliatsikas & Marinis, 2013b; Silva & Clahsen, 2008). These studies reveal L2ers' reduced

sensitivity to the morphological structure of complex forms relative to NSs; hence L2ers rely on whole-word storage more than on compositionality. That is, the L2 grammar suffers from a *compositionality deficiency* during the processing of morphologically complex forms.

A *compositionality deficiency* in the L2 grammar refers to a deficit in the rule-based capacity for composing *regular bare forms* with an *–ed* suffix (but not *irregulars*). For instance, in English, only regular verbs are composed via application of a morphological rule by which the *–ed* suffix is combined with regular stems, but irregular verbs have somewhat idiosyncratic mappings. Because of such a compositionality deficiency in the L2 grammar, L2ers process both regular and irregular past tense forms as single units. Consequently, L2ers do not differentiate between regular and irregular past tense forms during processing.

This compositionality, as an essential and universal rule-governed principle of the human language system, is found at different levels of language processing, including morpheme, word, phrase, and sentence levels (Newman, Ullman, Pancheva, Waligura, & Neville, 2007). Additionally, not only does this compositionality govern the production system, but it also governs the comprehension system in the L1 (Marslen-Wilson, 2007). Reviewing L1 processing research of language comprehension, Marslen-Wilson (2007) argued that the morphological structure of morphologically inflected forms might not "participate in language comprehension as whole forms, but rather as bearers of inflectional" and stem morphemes conveying semantic and syntactic information (p. 189).

By focusing on L2ers' compositionality ability, the current study examines the effect of token frequency (the occurrence of a given word form within a corpus) on the processing of regular past tense forms in a reading-based experiment. It compares two groups of adult L2ers (L1 Persian L2ers &

L1 Arabic L2ers) of English with a group of NSs. The purpose is to explore as to whether NSs and L2ers are sensitive to frequency effects in detecting *regular past tense forms of English* (henceforth, regulars) and whether L2ers' language processing system differs from the native system. The rationale underlying the use of frequency as a diagnostic for L2ers' compositionality ability is that the absence of the frequency effect reflects their compositionality ability, whereas its presence indicates their *in*capability in compositionality (more detail in Section 0).

Literature Review

Effects of frequency on morphologically inflected forms in L1 & L2

Since this study uses *speeded* acceptability judgment tasks, an overview of the impact of frequency in the previous *speeded production* research is given below. In a speeded production experiment, participants listen to a verb stem for which they should produce an inflected form as quickly as possible while their production latencies are measured. The current study appears to be similar to speeded production studies because (a) processing pressure is a common feature of both speeded production and speeded acceptability judgment tasks. (b) "In production, the rule route is always activated" (Clahsen, Hadler, & Weyerts, 2004, p. 705). Likewise, in sentential contexts here, regular past tense forms need to be recognized for their syntactic role for which the past tense rule should be activated.

In L1, several studies have exhibited frequency effects for irregularly inflected forms but not regularly inflected forms. For instance, several studies have found this distinct pattern for English past tense forms (Beck, 1997; Prado & Ullman, 2009; Prasada, Pinker, & Snyder, 1990) and German past participles (Clahsen, Hadler, & Weyerts, 2004; Fleischhauer & Clahsen, 2012). In an early article, Prasada Pinker and Snyder (1990) used a speeded

production task to study frequency effects on English past tense forms. They found statistically shorter production times for high- than low-frequency irregular past tense forms, but, for regulars, they found a trend for a reverse frequency effect or an anti-frequency effect (longer production times for high- than low-frequency regulars), though it was not significant. The authors interpreted the frequency effect for irregular forms as indicating full-form representations and access from the mental lexicon. As for the *in*significant anti-frequency effect (i.e., the anti-frequency trend) for regular forms, they suggested that the processing of regular past tense forms is not sensitive to full-form frequency and thereby is not dependent on full-form representations. Consistent with Prasada et al.'s findings, more recently, Prado and Ullman (2009) have reported shorter production times for high- than for low-frequency irregular past tense forms of English.

In contrast, this frequency effect was smaller for regular past tense ones due to specific item- and participant related factors. As the difference between these effects in irregulars and regulars was significant, they interpreted this contrast as evidence for a dual-system view in which irregulars are processed as full forms. Still particular item- and participant-level factors like frequency, imageability, and gender can lead to the whole-form processing of regulars, too.

Similar to previous findings on English, Clahsen *et al*. (2004) found the same pattern of frequency effects in German regular and irregular participle forms. Using a speeded production task, they compared an adult group and two age groups of children (mean ages: 5;3-7;9 & 11;0-12;8). For irregulars, the results indicated that the participants in all age groups produced high-frequency irregular participle forms faster than low-frequency ones. Conversely, for regulars, children produced low-frequency regular participle forms more quickly than high-frequency ones creating an anti-frequency

effect. A post-hoc analysis demonstrated that this anti-frequency effect was modulated as a function of slowed lexical retrieval because this effect was also found in a subgroup of slow adults. Clahsen et al. proposed that this anti-frequency effect might arise from a slowed lexical retrieval displaying probable individual differences in working memory capacities. Notwithstanding, they interpreted their findings in terms of the dual-mechanism models suggesting that the mental mechanisms and representations of morphological processing are similar in children and adults.

Following Clahsen et al. (2004), Fleischhauer and Clahsen (2012) investigated the effect of word-form frequency and working memory capacity on the processing of inflection. Using a speeded production task, they tested adults' and children's spoken production of German past participles (e.g., *ge-frag-t* 'asked'$_{[PART,\ regular]}$, *ge-schlaf-en*'slept'$_{[PART,\ irregular]}$). For irregularly inflected participles, all age groups displayed a robust advantage of high- over low-frequency forms. As for regulars, children (but not adults) exhibited a significant anti-frequency effect. Considering working memory scores as one factor, they found that adults with low working memory capacity also performed like children displaying a frequency effect for high-frequency irregulars and an anti-frequency effect for high-frequency regulars. Following Clahsen et al.'s (2004) study, Fleischhauer and Clahsen (2012) also interpreted their findings in terms of a dual mechanism account. However, they suggested that the low capacity of the working memory might be the reason for the anti-frequency effect in regulars. That is because this effect was found in children and a subgroup of adults with low working memory capacity.

In L2, a few studies have examined the frequency effects of inflectional morphology. For instance, using speeded production tasks for English past tense forms, Beck (1997) presented isolated words as stimuli to a group of

L2ers and NSs. The NSs revealed frequency effects for irregulars but not regulars. However, the L2ers did not display frequency effects on either verb type. Precisely, although, in experiments 3 (on NSs), 5 (on NSs) and 6 (on L2ers) high- and low-frequency regulars did not differ significantly in terms of production times, in experiments 1 (on NSs), 2 (on L2ers) and 4 (on L2ers) both groups displayed an anti-frequency effect. Overall, whereas the NSs (but not L2ers) demonstrated frequency effects for irregulars, both groups exhibited an anti-frequency effect for regulars. Thus, Beck's study did not indicate any such L1-L2 contrasts in the processing of high- and low-frequency regulars (see Lalleman et al., 1997 for similar findings). Unlike the findings in Beck's study, Babcock et al. (2012) found significant L1-L2 contrasts in the processing of English regular and irregular past tense forms. Using a production paradigm, they investigated regulars and irregulars with a mean surface form frequency of six words per million. They tested L2ers of English (Chinese & Spanish) and found similar frequency effects for both regulars and irregulars in both L2 groups. However, the NSs revealed a significant frequency effect only for irregulars. Babcock et al. argued that late L2ers over-rely on the storage of morphologically complex forms irrespective of their L1s (see also Bowden *et al*., 2010 for similar findings in L2ers of Spanish).

The anti-frequency effect during the processing of regulars in speeded production tasks has motivated three types of explanation. In an early attempt, Beck (1997) suggested that this effect occurs when the experimental list includes regular and irregular forms. In such a list, the phonological features common between certain irregulars and regulars may prevent the application of the morphological rule, creating some disturbance or an artifact. The evidence for this argument comes from her experiment 3, which included only regular forms, and the anti-frequency effect disappeared for NSs (but note that

L2ers continued to show the anti-frequency effect in experiment 4 with the same stimuli). Clahsen, Hadler, and Weyerts (2004) oppose this explanation arguing that Beck's experiment 3 involved only a very few distractors, which perhaps led participants to form a predictive strategy such that they might have predicted to produce regulars in the absence of irregulars in the experimental list. As a second proposal, several researchers maintain that the anti-frequency effect is likely due to slowed lexical access as a result of low working memory capacity in children and slow adults (Clahsen et al., 2004; Fleischhauer & Clahsen, 2012).

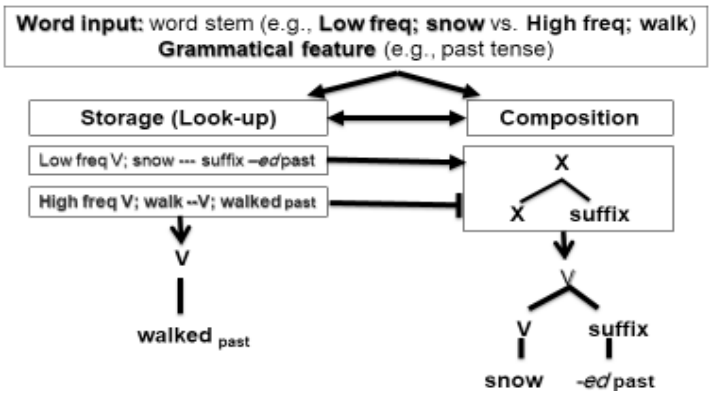

Figure 1: Simple representation of the Words and Rules theory (adapted from Pinker & Ullman, 2002)

As a third proposal, Pinker (1999) explains the anti-frequency effect in line with the dual-mechanism model (cf. **Figure 1**) in which two routes are assumed to be potentially accessible for the processing of morphologically complex forms, namely storage (lexical look-up) and composition (rule) routes. Pinker suggests that a more likely possibility is that for the production of the high-frequency regular forms, both storage and composition routes are

accessed in parallel. Accordingly, high-frequency regular forms (e.g., *walk*) may invoke memory traces via the storage route and activate the composition route simultaneously, as indicated with the bidirectional link between the storage and composition routes. Since the retrieval of high-frequency regulars from memory inhibits the composition route (demonstrated with the inhibitory link in **Figure 1**), it would take longer to produce high-frequency regulars that involve activation of both memory traces in the storage route and inhibition of the composition route. However, producing the low-frequency regulars (e.g., *snow*) would take less time because it only involves activating the composition route but not invoking memory traces. Evidence for people's creating memory traces for high-frequency regulars is provided by comprehension research using word vs. non-word lexical decision tasks in which regularly inflected forms showed frequency effects in several languages (Baayen, Dijkstra, & Schreuder, 1997; Sereno & Jongman, 1997; Taft, 1979). These observations may support the revised dual-mechanism models, which assumes that high- but not low-frequency regulars create memory traces in the mental lexicon (Alegre & Gordon, 1999; Pinker, 1999).

In sum, whereas previous speeded production studies on NSs and L2ers presented an almost clear picture of the effect of frequency on irregulars (a robust advantage for high- over low-frequency irregulars), the effect of frequency on regulars, is rather mixed and inconclusive for both NSs and L2ers in these studies.

L2 Morphological Processing Accounts

According to Silva and Clahsen (2008), the past tense debate has led to two broad accounts of L1-L2 processing contrasts, namely the shared systems hypothesis of Perani and Abutalebi (2005) and the Declarative-Procedural (DP) Model of Ullman (2005;2001). The shared systems hypothesis holds that

the mechanisms and neural substrates of morphological processing are essentially the same in the L1 and L2 processing. However, L2 processing is different in that it is more demanding than L1 processing and is likely influenced by L2ers' L1, age, proficiency, and cognitive factors such as working memory and processing speed (Perani & Abutalebi, 2005). For instance, McDonald (2006) holds that L2 processing is slower and more memory-demanding than L1 processing. In a speeded grammaticality judgment task comparing NSs and adult L2ers, she found that the NSs' performance under noise or memory stress was parallel to that of the adult L2ers without these stressors. That is, while the L2 processing is qualitatively similar to the L1 processing (it has the same mechanisms), it is quantitatively different and affected by low L2 working memory capacity, slow L2 processing speed, and reduced automaticity (for evidence from brain-imaging studies, see Indefrey, 2006; Perani & Abutalebi, 2005; Stowe & Sabourin, 2005; for evidence from ERP studies, see Friederici, 2002; Mueller, 2005; Rossi et al., 2006; Service et al., 2002; for evidence from behavioral studies, see Hoover & Dwivedi, 1988; Hopp, 2010; McDonald, 2006; for very recent evidence from behavioral and ERP studies, see Coughlin et al., 2019).

As an alternative view, akin to the dual-system theory (Pinker, 1999), Ullman's DP model (2005; 2001) holds that for L1 language processing, there are two routes each fed by two distinct memory systems. The lexical look-up route retrieves stored whole-form words (e.g., irregular past tense forms) from the declarative-memory system, and the rule-based system route computes grammatical rules, including morphological ones (e.g., *Add -ed* rule) in the procedural-memory system. For L2 processing, the DP model postulates that L2ers rely on the declarative memory system more than the procedural memory system, hence making no differentiation between past tense regulars and irregulars. However, Ullman (2005) speculates that, with increased

proficiency, L2ers might use the procedural memory system and process regular and irregular forms via the rule-based route and the lexical look-up route, respectively, as in L1 processing.

In line with the shared systems hypothesis, several recent L2 processing studies suggest that L2ers may rely on similar mechanisms to NSs when processing morphologically complex forms. For instance, by using a cross-modal priming lexical decision task, Basnight-Brown et al. (2007) found that similar to NSs, proficiency matched Serbian and Chinese late L2ers of English displayed significant priming effects in the processing of regular past tense form (e.g., guided-GUIDE, pushed-PUSH) (see Feldman et al., 2010, experiment 2, for similar findings in Serbian L2ers of English). Likewise, in a masked-priming lexical decision task in their experiment 1, Feldman et al. (2010) found that both NSs and high proficiency Serbian L2ers displayed priming effects for morphologically regular prime-target pairs (*billed–BILL*) but not for orthographically related prime and target pairs (*billion–BILL*). However, low proficiency L2ers exhibited priming effects for both the morphological and orthographic conditions. Even L2ers of Chinese, whose native language has no past tense feature,[2] were sensitive to compositionality in past tense forms. Using a masked-priming task, Dawei, Jinbo, and Heping (2016) found that high proficiency Chinese L2ers of English exhibited significant priming effects for morphologically prime-target pairs (*asked–ASK*) compared to morphologically related prime-target pairs (*ask–ASK*) and morphologically unrelated prime-target pairs (*live–ASK*). Moreover, L2 processing studies using other paradigms like Self-Paced Reading (SPR)[3] also

[2] Since Mandarin Chinese lacks surface forms indicating the Tense feature, Hawkins and Chan (1997) claim that the L2 grammar of L1 Chinese L2ers lacks the [past] Tense feature due to its absence in their L1.

3 In the SPR, participants read sentences word by word or phrase by phrase by pressing a key/button. Each button press measures RTs in milliseconds (ms). Each button press thus

found L1-L2 similarity. In an SPR task, Pliatsikas and Marinis (2013a) found that similar to NSs, highly proficient Greek L2ers of English displayed significantly longer latencies for regularly inflected past tense verbs than irregularly inflected verbs in grammatical sentences, irrespective of the type of L2 exposure (see also Hahne et al. (2006), for similar findings with highly proficient Russian L2ers of German via an Event-Related Potentials (ERPs) experiment). Very recently, Coughlin et al. (2019) have also found that not only high but also low proficiency L2ers of French exhibit native-like patterns in morphological priming in both behavioral and ERP experiments (see also Foote, 2015 who found that both intermediate and advanced L2ers of Spanish were sensitive to regularly inflected Spanish verbs in a native-like fashion). Summarizing, the above observations may support the shared systems hypothesis since both NSs and L2ers use whole-word look-up and the rule routes for the processing of irregulars and regulars, respectively.

In line with the DP model, several recent L2 processing studies suggest that L2ers rely more on the lexical look-up route than the rule-based route, and thus qualitatively become different from NSs during morphological processing. For instance, in a masked priming experiment, Silva and Clahsen (2008) found that, unlike NSs, advanced L2ers of English (Chinese, Japanese & German) did not display significant priming effects for morphological prime-target pairs (*prayed-PRAY*) relative to unrelated pairs (*baked-PRAY*). They concluded that irrespective of their native language, L2ers do not decompose morphologically complex words and instead access morphologically complex words as whole forms (see also a replication of this study in Clahsen et al., 2013 for similar results with advanced Arabic-speakers

provides insights into how fast or slow participants process each word (Just, Carpenter, & Wooley, 1982).

.

of L2 English, but see another replication of this study by Voga et al.**,** 2014 who found priming effects for English past tense regulars in Greek L2ers of English). Likewise, in a masked priming experiment, Neubauer and Clahsen (2009) found that, unlike NSs, advanced L1 Polish L2ers of German did not show a significant priming effect for regular past participles (*gespielt* '*played*'). However, both groups showed partial priming for irregular participles (*gelaufen* '*run*'). In their lexical decision experiment, the NSs showed shorter lexical decision times for high- than for low-frequency irregular forms and no frequency effect for regular ones. The L2ers, however, revealed similar frequency effects for both regular and irregular participles. Similarly, in a masked priming task, Pliatsikas and Marinis (2013b) tested regular (*Played-PLAY*) and irregular verbs (*kept-KEEP*) in NSs and L1 Greek L2ers of English and found priming effects for regular pairs only in NSs. Still, both groups displayed priming effects for irregular pairs. Summarizing, these observations may support the DP model. Based on this model, whereas NSs process regulars via the rule route and irregulars via the lexical look-up route, L2ers over-rely on the lexical look-up route to process both forms.

Altogether, the existing research on L2 morphological processing is rather mixed and inconclusive. Some of the observations may support the DP models indicating impairment in L2ers' compositionality ability and their over-reliance on the storage system. By contrast, some other studies are the opposite indicating that both NSs and L2ers process regulars in the composition system and irregulars in the storage system. So, these latter observations may support the shared systems hypothesis.

This Study

By using frequency as a diagnostic, the first objective of the current study was to add empirical evidence to the debate concerning the *storage* versus

composition of English regular past tense forms in L1 and L2 processing. The second objective was to explore as to whether L1 and L2 processing systems are fundamentally (i.e., qualitatively) the same or different. The underlying rationale for using frequency as a diagnostic is the fact that lexical items, which are more frequently available in the input due to their higher frequencies, are also accessible more easily from the mental lexicon (Prado & Ullman, 2009). Consequently, to reflect storage, regular forms of higher frequency should show frequency effects; they should be accessed more successfully than lower frequency forms. By contrast, to reflect composition, regular forms of lower frequency should lack frequency effects as they need a rule computation. Thus, in line with the objectives of the study, the following research questions are stated.

Question 1: Do participants (NSs or L2ers) activate the storage or composition route to detect high-frequency and low-frequency regulars?

Question 2: Do NSs and L2ers pattern alike or differently?

In accordance with the previous research, the revised dual-route models (Pinker, 1999), and the DP model (Ullman, 2005; 2001) the following predictions are tentative answers to the above questions. Note that Predictions 1 and 2 are relevant to Question 1 and Prediction 3 to Question 2.

Prediction 1: *If participants demonstrate frequency effects, they should detect high-frequency regulars with higher accuracy rates and shorter RTs than low-frequency ones. If so, this is indicative of their over-reliance on the storage route because high-frequency forms create stronger memory traces. As such, storage is used for high-frequency regulars.*

Prediction 2: *If they exhibit an anti-frequency effect, they should detect low-frequency regulars with higher accuracy rates and shorter RTs, compared to high-frequency regulars. If so, this is indicative of their over-reliance on the composition route. That is because low-frequency forms do not have memory traces, and thereby should be combinatorially made via the composition route. As such, composition is used for low-frequency ones.*

Prediction 3: If both groups demonstrate the same pattern of performance for the frequency effect, this can be evidence for the shared systems hypothesis (Perani & Abutalebi, 2005). Meanwhile, the DP model (Ullman, 2005; 2001) states that, at high proficiency levels, L2ers may over-rely on the procedural memory system (i.e., the composition route) for regular items (i.e., low-frequency ones in this study), and thus look like NSs.

To explore the above research questions, this study tested whether, compared to NSs, intermediate-to-advanced Persian and Arabic L2ers of English were sensitive to the frequency effect of regular past tense verbs in sentence-based contexts when they were under processing pressure. The participants' sensitivity was calculated by measuring the NSs' and L2ers' reactions (accuracy & Reaction Time (RT) data) to ungrammatical forms of high- and low-frequency regular past tense verbs compared to grammatical controls. The RT data in this task also provide processing pressure (see Section 0), which is also a potential factor giving rise to L2 variability (Ionin & Wexler, 2002; McDonald, 2006; Prévost & White, 2000).

Method

Speeded Acceptability Judgment (SAJ) Task

Speeded Acceptability Judgment (SAJ) tasks have widely been used to tap language processing in the L1 (McElree & Griffith, 1995; Schlesewsky & Frisch, 2003) and L2 sentence comprehension research (Clahsen, Felser, Sato, & Silva, 2010; Hopp, 2010; Sato & Felser, 2010). A SAJ task requires that participants react to the acceptability of stimuli as quickly and accurately as possible. Stimuli are broken into words or phrases and presented at a very high-speed rate via a Rapid Serial Visual Presentation (RSVP) mode. The RSVP mode presents words one at a time for a specified time interval (see **Figure 1**). Processing difficulty, defined as insensitivity, is shown by longer RTs and lower accuracy rates or both (McElree & Griffith, 1995).

Since the online SAJ task is useful for studies exploring implicit processing (Ellis, 2005), it is appropriate for the current research. Compared to the online SAJ task, in an offline task L2ers are notoriously reliant on meta-linguistically explicit knowledge when asked to judge sentences. However, this online task may potentially prevent them from relying on that explicit knowledge because the RSVP mode, which elicits forced responses, puts them under time pressure. Moreover, proponents of the Missing Surface Inflection Hypothesis[4] (Ionin & Wexler, 2002; Prévost & White, 2000) assert that L2ers show variability due to processing pressure. In that case, a SAJ task is preferable over a *Self*-Paced Reading task since, in the former, the L2ers are under processing pressure, whereas in the latter, they *themselves* control their pace of reading.

[4] This hypothesis posits that the absence of verbal inflection in L2ers'oral production, (when L2ers are under added time pressure) may indicate the absence of surface realization of inflection while the L2 grammar may not lack the relevant abstract syntactic features.

Participants

Seventeen intermediate-to-advanced L1 Persian speakers (8 females; 0 left-handed) and fifteen intermediate-to-advanced L1 Arabic speakers (10 females; 0 left-handed) were recruited from among students studying different subjects at the University of Essex. The Persian and Arabic L2ers were from among the Iranian, Saudi, and Kuwaiti student communities, respectively. All the participants were normal or corrected-to-normal in their vision and naïve regarding the purpose of the experiment. All the L2ers were residents in the UK at the time of testing, and on average, had been exposed to English for a mean of 2.6 in years for the Persian L2ers and a mean 1.67 in years for the Arabic ones. Initially, each group consisted of 20 participants, all of whom were screened for proficiency with the Oxford Placement Test (OPT) (Allan, 1992). Three from the Persian group and five from the Arabic group were removed due to their low proficiency scores. By equating the two L2 groups on proficiency, the Persian and Arabic L2ers were matched based on their proficiency scores [t (30) = -1.08, p=.276]. This equating resulted in a sample size of 32 for both L2 groups altogether. [5] The bio-data and their mean proficiency scores, as measured by the OPT, are given in

Table 1.

Table 1.
L2ers' Biodata and Oxford Placement Test (OPT)

	L1 Persian L2ers			L1 Arabic L2ers		
	Mean	SD	Range	Mean	SD	Range
Age (years)	28.45	5.3	23-40	28.9	6.8	23-46

[5] Both Persian and Arabic are similar in terms of some features potentially affecting L2ers' performance, namely they are both highly rich in their inflectional system, and have the same script and a right-to-left writing system.

Age of onset (years)	10.25	3	5-16	12.4	2.7	8-18
Length of residence (years)	2.6	2.7	0.11-10	1.67	1	0.5-4.6
OPT (total 100)	74.82	7	62-92	75.4	6.3	63-92

Note that the score scale of the OPT is for the overall score for the grammar section, with a maximum score of 100. All the remaining L2ers scored above 62, which corresponds to the intermediate-to-advanced level on the OPT scale. Therefore, they met the minimum requirement to participate in this study. 22 NSs (12 female, three left-handed, mean age: 25.5, range: 19 - 36) were also recruited from among the home students of the British English at the University of Essex to serve as controls. The three groups were provided with coffee and sweets, and the English NSs were paid £5 for their participation.

Materials

The participants made acceptability judgments for 120 sentences. Twenty were experimental items relevant for testing compositionality in past tense verbs. This group of items was further split into two groups depending on their token frequency (expressed as a logarithmic value, see

Appendix A) which is based on each verb's token frequency of occurrence per million in all genres (both written & spoken texts) in the Corpus of Contemporary American English (COCA) database (Davies, 2008): a high-frequency group (mean frequency= 1.72 per million) and a low-frequency (mean frequency = 0.34).

All items were simple active sentences each three words long. All regulars in the 20 sentences (20 grammatical & 20 ungrammatical) were intransitive. They were made ungrammatical by using bare unmarked forms (e.g., *Yesterday they cried* vs. **Yesterday they cry*). There were also filler items (N=100) testing grammatical and ungrammatical sentences of case, S-V Agreement, gender, animacy, and word order. Altogether, the list of items (experimental items and fillers) formed 120 trials. Each item appeared in two different conditions that were created by manipulating grammaticality (grammatical vs. ungrammatical). The total number of 120 grammatical and 120 ungrammatical items were arranged in two lists according to a Latin Square Design such that each participant saw each item in only one condition. All items were then pseudo-randomized such that experimental items did not occur consecutively (see the experimental list in

Appendix B).

Procedure

The participants first filled in a questionnaire to provide their bio-data (see

Table 1) and a consent form to participate in the current research. Both groups (NSs and L2ers) were tested in a sound-proof and quiet place (the Psycholinguistics Lab at the University of Essex).[6] The participants judged

[6] I myself collected the data for this study (as part of the requirements for an MA in psycholinguistics & neurolinguistics from the University of Essex).

the acceptability of the sentences presented to them on a 14-inch computer screen one word at a time. Each sentence trial started with a fixation cross displayed in the center of the screen for 500 milliseconds (ms) to signal that an item was about to appear. Afterward, a sentence was presented in the center of the computer screen in a word-by-word fashion at a rate of 350ms per word. Using the RSVP paradigm, the DMDX software (Forster & Forster, 2003) presented the words automatically replacing one another, as illustrated in **Figure 2** below. All words appeared in white letters on a black background in Arial Font of 30 points.

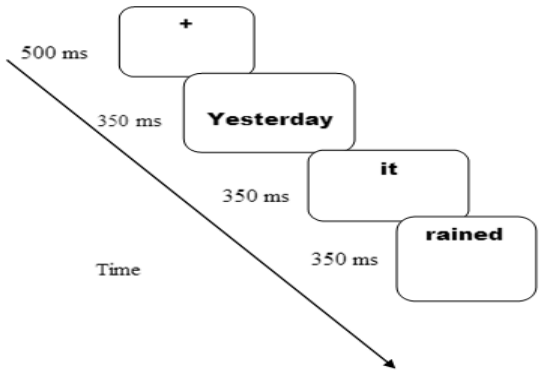

Figure 2. An illustration of the Rapid Serial Visual Presentation of items via DMDX.

The participants were required to read sentences and make acceptability judgments after reading the final word as quickly and accurately as possible. They used a game-pad by pressing either a '*Yes*' or '*No*' button to accept or to reject the sentences. The game-pad was activated from the onset of the final word. There were three breaks after each set of 40 items for which the participants were allowed to take a break if needed. Before the main experiment, a practice session comprising of 10 trials was run to let them be

familiarized with the task. The main experiment took approximately 35-45 minutes. The L2ers were also required to complete the grammar part of the OPT of English (Allan, 1992), which provided an index of their general proficiency.

Data Analysis

The analyses involved only the correct response trials in RT data. Individual outlier trials were trimmed from the data set. These outlier trials were defined as response times of an individual participant per condition, which fell above or below 2.5 standard deviations of the group mean. This affected 7 (1.59%) trials for the NSs and 24 (3.75%) for the L2ers.

For the statistical analysis, the data sets were analyzed using two types of mixed models. For the accuracy data, a Generalized Linear Mixed Model (GLMM) was used, which has a logistic link function and binomial variance for categorical dependent variables (e.g., accuracy rates =accurate vs. inaccurate) (Baayen, 2008; Jaeger, 2008). For the RT data, a Linear Mixed Model (LMM) was used since the dependent variable is continuous in RT data (Baayen, Davidson, & Bates, 2008). Mixed models provide a robust statistical method for analyzing experimental data (Pinheiro & Bates, 2000). Moreover, there is no need to consider the assumptions of homogenous variance and sphericity, which are inherent in ANOVAs (Pinheiro & Bates, 2000).

Mixed models were applied using R (R Development Core Team, 2017). Predictor variables were grand-mean centered to avoid issues of collinearity. Models were first fitted to the full data set for both groups; any interaction terms were further explored by analyzing data from each group separately. The model fitted to the data from the L2ers also tested whether proficiency was a significant predictor of their performance.

Following Barr et al. (2013), the initial full models consisted of all fixed factors and a maximal random structure. However, when any individual random variable reached a high correlation of +l or –l, it was removed from the maximal random structure. Fixed-effects were compared through contrasts between levels. Each level of a fixed factor was contrasted to a specified reference level shown in **bold type** below. The initial full model consisted of Group (**NSs** vs. L2ers**)**, Frequency (**High Freq** Vs. Low Freq), and Grammaticality (**GRAMM** vs. UNGRAMM) as fixed effects.

Results and Discussion

Results

In **Table 2**, an initial comparison between the Persian and Arabic groups ruled out any differences between the groups; only the negative coefficient for the significant main effect of Grammaticality indicates that accurate responses decreased in ungrammatical items relative to grammatical ones. Thus, these groups were combined to increase the power of the analyses.

Table 2.

Fixed-effects from GLMM fit to data from Arabic & Persian L2ers, Regular Past Tense (Accuracy Rates)

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.94134	0.19533	9.939	< 2e-16	***
Group (Arabic)	0.09839	0.25613	0.384	0.70086	
Frequency (Low)	-0.21202	0.37193	-0.57	0.56864	
Grammaticality (Ungrammatical)	-0.95467	0.34151	-2.795	0.00518	**
Group (Arabic) ×Frequency (Low)	0.31568	0.51292	0.615	0.53825	

Group (Arabic)× Grammaticality (Ungrammatical)	-0.17451	0.51157	-0.341	0.73301
Frequency (Low) × Grammaticality (Ungrammatical)	0.83536	0.67944	1.229	0.21889
Group (Arabic)× Frequency (Low) × Grammaticality (Ungrammatical)	-0.28276	1.0288	-0.275	0.78344
Formula in R: Accuracy~1 + Group *Frequency* Grammaticality +(1 |Item)+(1|Participant)				

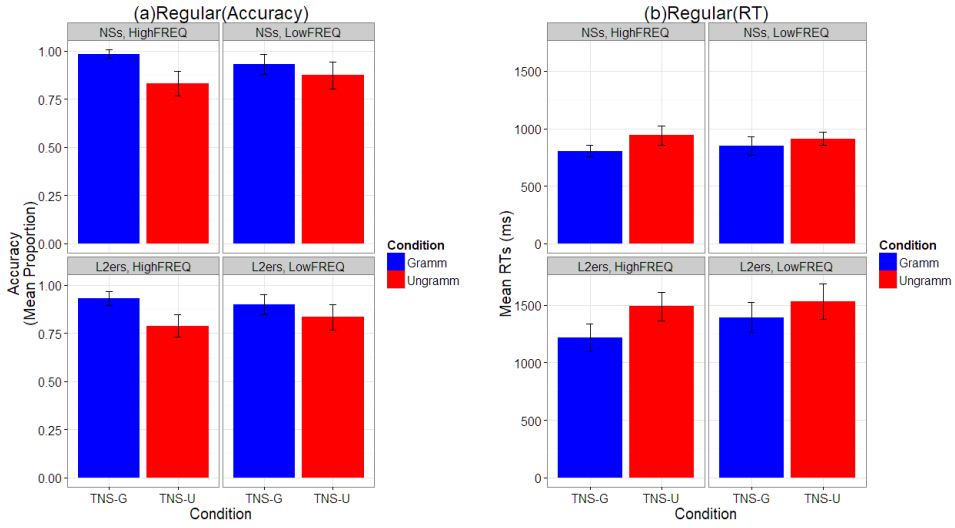

Figure 3. NSs vs. L2ers, Mean Accuracy (a) & RTs (b), Regular Past Tense (Error bars are 95% CI).

TNS-G: tense-grammatical
Yesterday we walked.

TNS-U: tense-ungrammatical
*Yesterday we walk.

Response Accuracy

In **Table 3,** the negative coefficient for the significant main effect of Group indicates that the NSs made more accurate responses than the L2ers. The negative coefficient for the main effect of Grammaticality demonstrates that accurate responses decreased in ungrammatical items relative to

grammatical ones. The significant two-way interaction (Frequency × Grammaticality) demonstrates between-frequency differences. Its positive coefficient indicates that accuracy rates increased when the verb frequency is low, and the sentence is ungrammatical (cf. Panel 'a' of **Figure 3**). This reverse effect is known as the anti-frequency effect. However, the three-way interaction (Group × Frequency ×Grammaticality) was not significant, indicating that L2ers displayed a native-like profile in detecting regular past tense verbs (cf. **Figure 3** (a)). In the bottom part of **Table 3**, the positive coefficient for the three-way interaction (Proficiency × Frequency × Grammaticality) was significant, indicating that as proficiency increased, L2ers' accuracy scores increased in detecting the ungrammatical low-frequency regulars compared to high-frequency ones. This is in line with the two-way interaction (Frequency × Grammaticality), indicating that both groups detected ungrammatical low-frequency regulars with higher accuracy scores than high-frequency ones.

Table 3.
Fixed-effects from GLMM fit to data from NSs & L2ers, Regular Past Tense (Accuracy Rates)

Fixed effects	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**	
(Intercept)	2.4297	0.2151	11.294	< 2e-16	***
Group (L2ers)	-0.7592	0.2798	-2.714	0.00665	**
Frequency (Low)	-0.3271	0.3513	-0.931	0.35185	
Grammaticality (Ungrammatical)	-1.4623	0.3089	-4.734	2.20e-06	***
Group (L2ers) ×Frequency (Low)	0.5699	0.5199	1.096	0.27299	
Group (L2ers)× Grammaticality (Ungrammatical)	0.8189	0.5576	1.469	0.14192	
Frequency (Low) × Grammaticality (Ungrammatical)	1.1563	0.4832	2.393	0.01672	*

Group (L2ers)× Frequency (Low) × Grammaticality (Ungrammatical)	-1.1854	1.0398	-1.14	0.25425	
Formula in R: Accuracy~1 + Group *Frequency* Grammaticality +(1 |Item)+(1 + Grammaticality |Participant)					
L2ers					
(Intercept)	2.06493	0.17064	12.101	< 2e-16	***
Proficiency	0.03861	0.02234	1.728	0.083907	.
Frequency (Low)	0.03656	0.32014	0.114	0.909075	
Grammaticality (Ungrammatical)	-0.95133	0.27135	-3.506	0.000455	***
Proficiency× Frequency (Low)	0.05683	0.04458	1.275	0.202341	
Proficiency × Grammaticality (Ungrammatical)	0.01665	0.04371	0.381	0.70327	
Frequency (Low) × Grammaticality (Ungrammatical)	1.03078	0.5525	1.866	0.062089	.
Proficiency× Frequency (Low) × Grammaticality (Ungrammatical)	0.22195	0.08975	2.473	0.013401	*
Formula in R: Accuracy~1 + Proficiency *Frequency*Grammaticality +(1 |Item)+(1 |Participant)					

Response Time (RT)

In the top part of **Table 4,** the positive coefficient for the significant main effect of Group indicates that the L2ers exhibited longer RTs than the NSs. The positive coefficient for the significant main effect of Grammaticality demonstrates that ungrammatical items were detected with longer RTs than the grammatical ones. The positive coefficient for the two-way interaction (Group × Frequency) was marginally significant, indicating that L2ers displayed longer RTs than NSs in low-frequency regulars than high-frequency ones. So further analyses are needed to explore the effect of frequency in each participant group. As is evident from the middle part of **Table 4**, for NSs the two-way interaction (Frequency × Grammaticality) was marginally significant. Its negative coefficient reflects the fact that NSs detected

ungrammatical low-frequency regulars with shorter RTs relative to the high-frequency ones, hence an anti-frequency effect. For L2ers, as is evident from the bottom part of **Table 4**, the two-way interaction (Frequency × Grammaticality) was not significant, but its negative coefficient indicates that they tended to process low-frequency regulars with shorter RTs relative to high-frequency ones; hence this is a trend, indicating the anti-frequency effect in L2ers. Note that this trend was also observed in other studies (Prado & Ullman, 2009; Prasada, Pinker, & Snyder, 1990). Additionally, the two-way interaction (Proficiency × Grammaticality) was significant. Its negative coefficient reflects the fact that relative to moderate proficiency L2ers, high proficiency ones detected ungrammatical regulars with shorter RTs.

Table 4.

Fixed-effects from LMM fit to data from NSs & L2ers, Regular Past Tense (RTs)

Fixed effects:	**Estimate**	**Std. Error**	**Df**	**t value**	**Pr(>|t|)**	
(Intercept)	1198.71	57.59	23.9	20.816	< 2e-16	***
Group (L2ers)	541.94	39.45	924.2	13.739	< 2e-16	***
Frequency (Low)	62.3	50.04	18.3	1.245	0.2288	
Grammaticality (Ungrammatical)	173.02	37.69	907.5	4.59	5.05e-06	***
Group (L2ers) ×Frequency (Low)	129.79	77.88	907.3	1.667	0.0959	.
Group (L2ers)× Grammaticality (Ungrammatical)	91.84	76.55	910	1.2	0.2305	
Frequency (Low) × Grammaticality (Ungrammatical)	-124.76	78.32	915.1	-1.593	0.1115	
Group (L2ers)× Frequency (Low) × Grammaticality (Ungrammatical)	-32.93	156.04	909.2	-0.211	0.8329	
Formula in R: RT~1 + Group * Frequency*Grammaticality +(1 |Item)+(1 |Participant)						

NSs						
(Intercept)	879.4183	47.73534	26.2	18.423	< 2e-16	***
Frequency (Low)	0.06218	46.33534	17.6	0.001	0.998944	
Grammaticality (Ungrammatical)	108.8734	27.96271	359.1	3.894	0.000118	***
Frequency (Low) × Grammaticality (Ungrammatical)	-101.784	57.70119	360.1	-1.764	0.078583	.
Formula in R: RT~1 + Frequency *Grammaticality +(1 |Item)+(1 |Participant)						
L2ers						
(Intercept)	1439.511	104.666	21.7	13.753	3.33e-12	***
Proficiency	4.634	6.859	365.3	0.676	0.499735	
Frequency (Low)	117.729	87.125	18.2	1.351	0.193148	
Grammaticality (Ungrammatical)	204.481	55.38	509.4	3.692	0.000246	***
Proficiency× Frequency (Low)	6.109	8.452	508.4	0.723	0.470175	
Proficiency × Grammaticality (Ungrammatical)	-26	8.326	517.2	-3.123	0.001892	**
Frequency (Low) × Grammaticality (Ungrammatical)	-126.03	115.532	513.4	-1.091	0.27584	
Proficiency× Frequency (Low) × Grammaticality (Ungrammatical)	7.707	17.227	520.4	0.447	0.654799	
Formula in R: RT~1 + Proficiency * Frequency *Grammaticality +(1 |Item)+(1 |Participant)						

Additional Analysis of Accuracy Data

To explore whether the observed anti-frequency effect was influenced by the individuals' potential differences in their speed of lexical processing, participants in each group were split into two subgroups ('fast' & 'slow') according to the median of the overall response latencies for their accurate

responses to the two experimental conditions (high- & low- frequency regulars) (NSs' median RT= 767.35 & L2ers' median RT= 1184.34). Separate GLMMs were conducted on the accuracy data for each language group with the factors RT Group (**Fast RT** vs. Slow RT), Frequency (**High Freq** Vs. Low Freq), and Grammaticality (**GRAMM** Vs. UNGRAMM) as fixed effects. As **Table 5** demonstrates, the separate GLMMs fitted to the accuracy data set from both NSs and L2ers did not yield any significant main effect of frequency or any interaction between frequency and RT Group in both groups of NSs and L2ers (*Fs*<1). Summarizing, the anti-frequency effect for regulars was not found in the subgroups of both groups indicating that slowed lexical access was not likely to affect the observed anti-frequency effect.

Table 5.

Fixed-effects from separate GLMMs fit data from RT Groups of NSs & Persian L2ers, Regular Past Tense (Accuracy Rates)

NSs					
Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	3.3389	0.4868	6.859	6.94e-12	***
RT Group (Slow)	0.6329	0.6016	1.052	0.2928	
Frequency (Low)	-0.6135	0.6995	-0.877	0.3804	
Grammaticality (Ungrammatical)	-2.2188	0.5701	-3.892	9.94e-05	***
RT Group (Slow) ×Frequency (Low)	-0.3222	1.0349	-0.311	0.7555	
RT Group (Slow) × Grammaticality (Ungrammatical)	2.0176	1.0834	1.862	0.0626	.
Frequency (Low) × Grammaticality (Ungrammatical)	1.8579	1.009	1.841	0.0656	.

RT Group (Slow)× Frequency (Low) × Grammaticality (Ungrammatical)	0.4998	2.026	0.247	0.8052	
Formula in R: Accuracy~1 + RT Group *Frequency* Grammaticality +(1 |Item)+(1 |Participant)					
L2ers					
(Intercept)	2.11777	0.17475	12.119	< 2e-16	***
Proficiency	0.05042	0.02408	2.094	0.036264	*
RT Group (Slow)	0.47057	0.3041	1.547	0.121763	
Frequency (Low)	0.11648	0.34905	0.334	0.7386	
Grammaticality (Ungrammatical)	-1.03692	0.30174	-3.436	0.000589	***
Proficiency× RT Group (Slow)	0.04995	0.04845	1.031	0.302557	
Proficiency× Frequency (Low)	0.08649	0.05084	1.701	0.088873	.
RT Group (Slow) × Frequency (Low)	0.84967	0.62736	1.354	0.175621	
Proficiency × Grammaticality (Ungrammatical)	0.01301	0.04815	0.27	0.787022	
RT Group (Slow) × Grammaticality (Ungrammatical)	0.53723	0.60726	0.885	0.376336	
Frequency (Low) × Grammaticality (Ungrammatical)	1.0953	0.62202	1.761	0.078258	.
Proficiency× RT Group (Slow) ×Frequency (Low)	0.05589	0.10227	0.547	0.584713	
Proficiency× RT Group (Slow) × Grammaticality (Ungrammatical)	-0.12175	0.09708	-1.254	0.209813	
Proficiency× Frequency (Low) × Grammaticality (Ungrammatical)	0.22596	0.10175	2.221	0.026368	*
RT Group (Slow) × Frequency (Low) × Grammaticality (Ungrammatical)	-0.77051	1.25483	-0.614	0.539191	

Proficiency ×RT Group (Slow) × Frequency (Low) × Grammaticality (Ungrammatical)	0.18255	0.20488	0.891	0.372921
Formula in R: Accuracy~1 +Proficiency*RT Group*Frequency*Grammaticality +(1 |Item)+(1|Participant)				

Summary of Results

This study aimed to test the effect of frequency on regular verb processing with two questions, and three predictions as follows: (1) Do participants (NSs or L2ers) activate the storage or composition route to detect high-frequency and low-frequency regulars? (2) Do NSs and L2ers pattern alike or differently? Predictions 1 & 2 are tentative answers to question 1, and Prediction 3 to question 2.

In general, despite NSs' faster RTs and higher accuracy scores, both groups exhibited the same pattern as evidenced by the lack of interaction. The results are as follows.

Accuracy Data

- Both NSs and L2ers displayed anti-frequency effects; they detected ungrammatical low-frequency regulars with higher accuracy scores, compared to high-frequency ones.
- Proficiency increased accuracy scores in low-frequency regulars compared to high-frequency ones, displaying a larger anti-frequency effect for higher proficiency L2ers (i.e., high proficiency L2ers obtained higher accuracy scores in low-frequency regulars than in high-frequency ones).
- Additional analysis of the accuracy data for both participant groups demonstrated that slowed lexical access was not the likely reason for the observed anti-frequency effect in the accuracy data.

RT data

- NSs detected ungrammatical low-frequency regulars with shorter RTs relative to the high-frequency ones leading to an anti-frequency effect, though this effect was marginal. L2ers also exhibited an anti-frequency *trend* though it was not significant.
- Relative to moderate proficiency L2ers, high proficiency ones detected ungrammatical regulars with shorter RTs.

Accordingly, the observed anti-frequency effect and the L1-L2 similar performance profile are consistent with **Predictions 2** and **3** but not **Prediction 1**.

General Discussion

By comparing with the previous research, this section focuses on three essential findings from the current study: (1) the presence of the anti-frequency effect, (2) the convergent pattern between the NSs and L2ers in accuracy and RT data, and (3) the proficiency effect.

Anti-frequency Effect

Both the NSs and L2ers exhibited an anti-frequency effect in both their accuracy and RT data indicating a significant disadvantage for high- over low-frequency regulars. This anti-frequency effect replicates previous speeded production studies on regulars for both child and adult NSs and L2ers (e.g., Beck, 1997, experiments 1, for adult NSs & 2, for L2ers & 4, for L2ers of English; Clahsen, Hadler, & Weyerts, 2004, for adult & child NSs of German; Fleischhauer & Clahsen, 2012, for adult & child NSs of German).

How can this anti-frequency effect be explained? Beck (1997) attributes this effect to the presence of an experimental artifact. She held that with a large number of irregulars in the experimental list, participants were biased

towards accessing stored lexical entries for every trial, hence their tendency for an anti-frequency effect. However, in the experimental list of the current study, first, irregulars were absent. So, the absence of irregulars might not bias the participants towards stored lexical entries. Second, instead of word lists, regulars were presented in sentential contexts (but not in isolation), and generally, sentences are not normally stored. Accordingly, the observed anti-frequency effect may not be due to an experimental artifact, as hypothesized by Beck.

As a second proposal, several researchers attributed such an anti-frequency effect to slowed lexical access due to the low working memory capacity in slow adults and children (Clahsen, Hadler, & Weyerts, 2004; Fleischhauer & Clahsen, 2012). However, in the current study, the additional analysis of the accuracy data revealed that the anti-frequency effect was not influenced by the slowed lexical retrieval because both slow and fast RT groups demonstrated such an anti-frequency effect. Furthermore, it seems unlikely that the detection of acceptability in such sentences (e.g., *Yesterday they cried* vs. **Yesterday they cry*) taxes L2ers' memory, let alone NSs (see Shibuya & Wakabayashi, 2008) for counter-evidence). This is because, in the present study, there was no need to check features as they are required for dependencies like Subject-Verb agreement. Hence, slowed lexical retrieval due to the low working memory capacity might not fully account for the observed anti-frequency effect in both groups.

As a third proposal, the anti-frequency effect seems to be consistent with the revised dual mechanism model (Pinker, 1999), which claims that word-form frequency affects high-frequency regulars reflecting parallel processing of retrieving whole-word forms and activating the rule-based system simultaneously. Eventually, as the storage route finds more memory traces for high-frequency regulars it blocks the composition route; hence high-

frequency regulars are retrieved via the storage route with lower accuracy rates and longer RTs. By contrast, low-frequency regulars are composed in the composition route, which is not affected by frequency effects, and thus leads to higher accuracy rates and shorter RTs. Accordingly, the presence of the anti-frequency effect demonstrates that both routes of lexical access were activated in the participants' mental lexicon: the storage route for high-frequency regulars and the composition route for the low-frequency regulars. Since the current study used the SAJ task in which the participants were under processing pressure, it is likely that both NSs and L2ers displayed such an anti-frequency effect for regulars due to the added processing pressure. This is because this anti-frequency effect was also found in *speeded* production studies with potentially similar processing pressure, as in the *speeded* acceptability task used in the current study (cf. Section 0.) See also McDonald's (2006) study, in which both NSs and L2ers performed similarly worse in detecting grammaticality of sentences under processing pressure. Thus, this anti-frequency effect may better be captured in terms of the revised dual mechanism model (Pinker, 1999), perhaps when participants are under processing pressure.

Convergent Pattern Between NSs vs. L2ers

Even though L2ers were not as good as NSs in their accuracy scores and RTs, their convergent pattern in the accuracy scores and RTs may reveal the fact that both NSs and L2ers retrieve high- and low-frequency regulars in a similar way. This may reflect the fact that the L2ers differ from the NSs quantitatively but not qualitatively.

Qualitatively, both NSs and L2ers displayed the anti-frequency effect indicating that they were similar in using both routes of lexical access: the storage and the composition routes for high-frequency and low-frequency

regulars, respectively. Therefore, the presence of the anti-frequency effect for both groups can provide evidence against studies (Babcock, Stowe, Maloof, Brovetto, & Ullman, 2012; Bowden, Gelfand, Sanz, & Ullman, 2010; Clahsen, Balkhair, Schutter, & Cunnings, 2013; Jacob, Hever, & Veríssimo, 2018; Neubauer & Clahsen, 2009; Pliatsikas & Marinis, 2013b; Silva & Clahsen, 2008) claiming that even advanced L2ers rely on direct whole-word access for both decomposable (e.g., *regulars*) and indivisible (e.g., *irregulars*) morphemes. Instead, the current study supports studies (Basnight-Brown et al., 2007; Beck, 1997; Birdsong & Flege, 2001; Coughlin et al., 2019; Feldman, Kostic, & Basnight-Brown, 2010; Hahne, Mueller, & Clahsen, 2006; Lalleman, van Santen, & van Heuven, 1997; Pliatsikas & Marinis, 2013a; Voga, Anastassiadis-Symeonidis, & Giraudo, 2014) which claim that L2ers are like NSs in activating both routes of lexical access. Accordingly, this study contributes new evidence from the accuracy and RT data claiming that both NSs and L2ers apply similar representations and processing mechanisms to detect inflected word forms. The presence of the anti-frequency effect in the SAJ task of the current study seems to be unprecedented because, so far, the anti-frequency effect has often been reported in speeded production studies.

Quantitatively, however, their different performance in RT data indicates that the L2ers are somewhat slower than NSs. That is, whereas the L2 processing does not differ qualitatively from the L1 processing, it differs quantitatively due to the potential low L2 working memory capacity, slow L2 processing speed, and reduced automaticity. In particular, very recently reviewing some psycholinguistic research of the L2 processing, Hopp (2018) argued that a slower time-course of lexical access might lead to the attenuated or delayed structure building in sentence processing (see Hopp, 2018 for more detail). Bearing this in mind, in the current study, slower lexical access in L2

processing might have led to a quantitatively different pattern from the NSs in RT data. Thus, in accordance with the recent findings (Coughlin, Fiorentino, Royle, & Steinhauer, 2019; Hopp, 2010; Indefrey, 2006; McDonald, 2006; Perani & Abutalebi, 2005; Rossi, Gugler, Friederici, & Hahne, 2006; Stowe & Sabourin, 2005), the above observations may support the shared systems hypothesis. This hypothesis claims that both NSs and L2ers use the same mechanisms during processing and that the difference between the NSs and the L2ers is quantitative rather than qualitative. Yet, L2ers' performance is slower and more error-prone than that of the NSs, perhaps due to a potentially slower time of lexical access compared to NSs.[7]

Proficiency Effect

As was discussed in Section 0, Ullman (2005) speculates that proficiency can help the shift from the use of the declarative to the procedural memory system in L2ers. Consistent with this point, in the current study, highly proficient L2ers exhibited such dissociation in their accuracy data. Relative to moderate proficiency L2ers, high proficiency ones displayed higher accuracy scores in low-frequency regulars compared to high-frequency ones leading to a larger anti-frequency effect for higher proficiency L2ers. This pattern of higher accuracy scores in low-frequency regulars than in high-frequency ones might mean that high proficiency L2ers automatized the *Add–ed* rule in their procedural memory system, even though these regulars had low frequency.

Moreover, for RT data, relative to moderate proficiency L2ers, high proficiency ones detected ungrammatical regulars with shorter latencies.

[7] Note that, in Section 0, the presence of the anti-frequency effect, as a common feature in both NSs and L2ers' performance, was attributed to the effect of processing pressure on both groups' performance in general, as in McDonald's (2006) study. However, in Section 0, slower lexical access in L2ers is likely to be the potential factor, which distinguishes NSs from L2ers since L2ers usually show longer RTs than NSs.

Consistent with this observation, proficiency was also reported to play an important role in the native-like performance in some L2 studies on subject-verb agreement (e.g., Ojima, Nakata, & Kakigi, 2005; Rossi *et al*., 2006; Safaie, 2015). As for the past tense of L2 English, Dawei, Jinbo, and Heping (2016) found the effect of proficiency on L2ers' processing of past tense inflection. Pliatsikas and Marinis (2013a) found no correlation between L2ers' proficiency levels and their mean RT and accuracy scores for the processing of regular past tense forms. Yet, they argued that L2ers at high proficiency levels performed at the ceiling, and thus they automatized the past tense rule. Similarly, very recently, Coughlin *et al*. (2019) have found that both low and high proficiency L2ers of French are sensitive to verb inflections (see also Foote, 2015 for similar findings with intermediate and advanced L2ers of Spanish).

Related to the effect of proficiency is the effect of the formal instruction on L2ers' compositionality ability as a possible explanation for their native-like performance. Since L2ers learned L2 English in the classroom, explicit grammatical instruction is likely to affect their proper use of regular past tense forms. This very explicit instruction may thus contribute to L2ers' native-like compositionality ability. However, this possibility may not explain the native-like behaviour. To begin with, given the effect of explicit instruction on L2ers' native-like performance, why proficiency should lead to a reverse effect on frequency (i.e., compared to moderate proficiency L2ers, high proficiency L2ers' accuracy scores decreased in high-frequency regulars compared to low-frequency ones). That is, explicit instruction may not lead to the difference between high proficiency and moderate proficiency L2ers. If it had had such an effect, high proficiency L2ers should have gained more accuracy scores in high-frequency regulars compared to low-frequency ones, but they did not. Moreover, explicit instruction should not lead to the significant

difference between the high-frequency and low-frequency regulars in terms of the accuracy rates, because after all, both forms are made with the same explicit morphological rule. Thus, this native-like behaviour is less likely to be attributed to the effect of explicit instruction on L2ers' compositionality ability rather, as discussed above, it may be that L1 and L2 processing mechanisms are similar and the L1-L2 difference is quantitative but not qualitative.

Conclusion & Pedagogical Implication

Generally speaking, despite NSs' faster RTs and higher accuracy scores, both NSs and L2ers demonstrated the same pattern of accuracy rates and RTs. In particular, the results of the current study are as follows. First, for the accuracy data, word-form frequency created an anti-frequency effect in both NS and L2er groups displaying higher accuracy rates for low-frequency regulars than for high-frequency ones. Second, proficiency increased L2ers' accuracy scores in low-frequency regulars compared to high-frequency ones displaying a larger anti-frequency effect for higher proficiency L2ers. As for the RT data, first, while the NSs exhibited a marginally anti-frequency effect, L2ers displayed a nonsignificant *trend* for the anti-frequency effect. Second, proficiency decreased L2ers' RTs in detecting ungrammatical verbs (i.e., relative to moderate proficiency L2ers, high proficiency ones detected ungrammatical regulars with shorter RTs).

These findings may provide evidence for the revised dual-mechanism models (Pinker, 1999) according to which accessing high-frequency regulars requires simultaneous activation of the composition and storage routes, first, and then blocking the composition route. Accordingly, when the high-frequency regulars are blocked by the composition route, they have to be accessed via the storage route, which leads to an anti-frequency effect. By

contrast, low-frequency regulars are accessed via the composition route per se without activating the storage route. Thus, the anti-frequency effect may reflect the fact that L2ers were efficient in using both routes of lexical access, namely the composition route for low-frequency regulars and the storage route for high-frequency ones. Additionally, since both NSs and L2ers followed the same pattern in activating both routes for the detection of high-frequency (in the storage route) and low-frequency regulars (in the composition route), the findings may support the shared systems hypothesis (Perani & Abutalebi, 2005). This hypothesis claims that the mechanisms and neural substrates of morphological processing are essentially the same in the L1 and L2 processing. Moreover, considering the role of proficiency, these findings also support the Declarative-Procedural Model (Ullman, 2005) claiming that proficiency can lead to the use of the native-like procedural memory system in L2ers. It was also suggested that the anti-frequency effect in regulars for both NSs and L2ers is possibly due to the effect of processing pressure in SAJ and speeded production tasks.

As for the pedagogical implication, the absence of frequency effect (i.e., anti-frequency) in the processing of morphologically complex forms is a diagnostic of L2ers' compositionality ability. Accordingly, compositionality, as a universal rule-governed principle of the human language system (Newman et al., 2007), could be considered as a universal principle in line with Universal Grammar (UG). According to this assumption (and following a UG-based perspective to second language acquisition like White, 2003), compositionality in the sense of the application of an inflectional rule may not need to be focused on in classroom contexts as such. Perhaps, only a little input is sufficient to trigger this ability in L2ers' grammar since low-frequency (but not high-frequency) regulars were detected by the composition route in the current study. That is because compositionality, as part of UG, has already

been instantiated innately before L2ers start learning a second language. This implication is based on the following reasons. First, in the current study, L2ers over-relied on the composition route to detect low-frequency regulars; hence their L2 grammar does not suffer from the compositionality deficiency. Second, explicit instruction might not account for L2ers' compositionality ability because high proficiency L2ers displayed a larger anti-frequency effect (i.e., high proficiency L2ers obtained higher accuracy scores in low-frequency regulars than in high-frequency ones) (cf. Section 0). If explicit instruction had increased L2ers' compositionality ability, it should have led to better performance in high-frequency than low-frequency regulars but it did not. As a conclusion, because compositionality has already been represented in the L2 grammar, it may not require to be focused on in classroom contexts; only a little input is sufficient to activate this ability in the L2 grammar.

References

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language, 40*, 41–61.

Allan, D. (1992). The Oxford Placement Test (2nd Edition). Oxford: Oxford University Press.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A practical introduction to statistics using R.* Cambridge, UK: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: evidence for a parallel dual-route model. *Journal of Memory and Language, 37*, 94–117.

Babcock, L., Stowe, J. C., Maloof, C. J., Brovetto, C., & Ullman, M. T. (2012). The storage and composition of inflected forms in adult-learned second language: A study of the influence of length of residence, age of arrival, sex, and other factors. *Bilingualism: Language and Cognition, 15*, 820–840.

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*, 255–278.

Basnight-Brown, D. M., Chen, L., Hua, S., Kostić, A., & Feldmann, L. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language, 57*, 65 – 80.

Beck, M. L. (1997). Regular verbs, past tense and frequency: tracking down a potential source of NSs/NNS competence differences. *Second Language Research, 13*, 93-115.

Birdsong, D., & Flege, J. (2001). Regular-irregular dissociations in L2 acquisition of English morphology. In *BUCLD 25: Proceedings of the 25 Annual Boston University Conference on Language Development* (pp. 123-132). Boston, MA: Cascadilla Press.

Bowden, H. W., Gelfand, M. P., Sanz, C., & Ullman, M. T. (2010). Verbal inflectional morphology in L1 and L2 Spanish: A frequency effects study examining storage versus composition. *Language Learning, 60*, 44–87.

Clahsen, H., & Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua, 120*, 2627–2637.

Clahsen, H., Balkhair, L., Schutter, J. S., & Cunnings, I. (2013). The time course of morphological processing in a second language. *Second Language Research, 29*(1), 7–31.

Clahsen, H., Felser, C., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning, 60*, 21-43.

Clahsen, H., Hadler, M., & Weyerts, H. (2004). Speeded production of inflected words in children and adults. *Journal of Child Language, 31*, 683-712.

Coughlin, C., Fiorentino, R., Royle, P., & Steinhauer, K. (2019). Sensitivity to inflectional morphology in a non-native language: evidence from ERPs. *Frontiers in Communication, 4*(21), 1-14. doi:10.3389/fcomm.2019.00021

Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990-present. Retrieved from Available online at http://corpus.byu.edu/coca/

Dawei, W., Jinbo, X., & Heping, W. (2016). The English past tense debate in the Chinese EFL learners' mind: A masked priming study. *Chinese Journal of Applied Linguistics, 39*(2), 185-198. doi:10.1515/cjal-2016-0012

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition, 10*, 101–116.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psycholinguistic study. *Studies in Second Language Acquisition, 27*, 141-172.

Feldman, L. B., Kostic, A., & Basnight-Brown, D. D. (2010). Morphological facilitation for regular and irregular verb formations in native and non-native speakers: little evidence for two distinct mechanisms. *Bilingualism: Language and Cognition, 13*, 99-118.

Fleischhauer, E., & Clahsen, H. (2012). Generating inflected word forms in real time: evaluating the role of age, frequency, and working memory. In A. Biller, E. Chung, & A. Kimball (Ed.), *Proceedings of the Annual*

Boston University Conference on Language Development. Vol. I, pp. 164-76. Somerville, MA: Cascadilla Press.

Foote, R. (2015). The storage and processing of morphologically complex words in L2 Spanish. *Studies in Second Language Acquisition, 39*, 735–767. doi:10.1017/S0272263115000376

Forster, K., & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers, 35*(1), 116-124.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In. In J. Altarriba, & R. Herridia (Eds.), *Bilingual sentence processing* (pp. 218–236). Amsterdam: Elsevier.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science, 6*, 78–84.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30*, 251–266.

Hahne, A., Mueller, J., & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience, 18*(1), 121-134.

Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional feature hypothesis'. *Second Language Research, 13*, 187-226.

Hoover, M. L., & Dwivedi, V. D. (1988). Syntactic processing in skilled bilinguals. *Language Learning, 48*, 1–29.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua, 120*, 901–931.

Hopp, H. (2018). The Bilingual Mental Lexicon in L2 Sentence Processing. *Second language, 17*, 5-27.

Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: which suggested differences can we trust and what do they mean? *Language Learning, 56*(Suppl. 1), 279–304.

Ionin, T., & Wexler, K. (2002). Why is 'is' easier than '-s'?: Acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research, 18*(2), 95-136.

Jacob, G., Hever, V., & Veríssimo, J. (2018). 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. doi:10.1177/1367006916688333

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446.

Just, M. A., Carpenter, P. A., & Wooley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*(2), 228-238.

Lalleman, J. A., van Santen, A. J., & van Heuven, V. (1997). L2 processing of Dutch regular and irregular verbs. *Review of Applied Linguistics, 115/116*, 1–26.

Marslen-Wilson, W. D. (2007). Morphological processes in language comprehension. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 175 –193). Oxford: Oxford University Press.

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*, 381-401.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 134-157.

Morgan Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language learning, 60*, 154–193.

Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*, 121-151.

Neubauer, K., & Clahsen, H. (2009). Decomposition of inflected words in a second language. *Studies in Second Language Acquisition, 31*, 403–435.

Newman, A. J., Ullman, M. T., Pancheva, R., Waligura, D. L., & Neville, H. J. (2007). An ERP study of regular and irregular English past tense inflection. *NeuroImage, 34*, 435–445.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience, 17*, 1212-1228.

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current opinion in neurobiology, 15*(2), 202-206.

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS statistics and computing*. New York: Springer.

Pinker, S. (1999). *Words and rules: The ingredients of grammar*. New York: Basic Books.

Pliatsikas, C., & Marinis, T. (2013a). Processing of regular and irregular past tense morphology in highly proficient second language learners of English: A self-paced reading study. *Applied Psycholinguistics, 34*, 943-970.

Pliatsikas, C., & Marinis, T. (2013b). *Rule-based morphological processing in a second language: a behavioural investigation.* The University of Reading. Retrieved from http://www.reading.ac.uk/english-language-and-appl...

Prado, E., & Ullman, M. T. (2009). Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology: Learning Memory, and Cognition, 35*, 849-866.

Prasada, S., Pinker, S., & Snyder, W. (1990). Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. *Paper presented at the 31st Annual Meeting of the Psychonomic Society.* New Orleans, LA.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16*, 103-33.

R Development Core Team. (2017). *R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria*. Retrieved from http://www.R-project.org

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 18*, 2030-2048.

Safaie, E. (2015). *Sensitivity to subject-verb agreement in second language sentence processing: evidence from L1 Persian speakers of L2 English.* The University of Essex.

Sato, M., & Felser, C. (2010). Sensitivity to morphosyntactic violations in English as a second language. *Second language, 9*, 101-118.

Schlesewsky, M., & Frisch, S. (2003). Nominative case as a multidimentional default. In S. Fischer, R. Vijvar , & R. Vogel (Eds.), *Experimental Studies in Linguistics 1, Linguistics in Potsdam* (Vol. 21, pp. 61-72).

Sereno, J., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition, 25*, 425–37.

Service, E., Simola, M., Metsanheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology, 14*, 383–408.

Shibuya, M., & Wakabayashi, S. (2008). Why are L2 learners not always sensitive to subject-verb agreement? In L. Roberts, M. Florence, & D. Annabelle (Eds.). John Benjamins Publishing Company.

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition, 11*(02), 245-260.

Stowe, L. A., & Sabourin, L. (2005). Imaging the processing of a second language: effects of maturation and proficiency on the neural processes involved. *International Review of Applied Linguistics, 43*, 329–353.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition, 7*, 263–72.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*(1), 105-122.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model. In C. Sanz (Ed.), *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice* (pp. 141-178). Washington, DC: Georgetown University Press.

Voga, M., Anastassiadis-Symeonidis, A., & Giraudo, H. (2014). Does morphology play a role in L2 processing? Two masked priming experiments with Greek speakers of ESL. *Lingvisticae Investigationes*(37), 338–352.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge: Cambridge Press.

Appendix A: List of regulars with token frequencies from COCA corpus

REGULAR	Length	Freq. PM	Log(10) PM
Quarrelled	10	0.53	-0.28
Snowed	6	0.88	-0.06
Moaned	6	2.47	0.39
Migrated	8	2.49	0.4
Giggled	7	2.95	0.47
Rained	6	2.98	0.47
Shaved	6	3.97	0.6
chuckled	8	5.31	0.73
Frowned	7	10.33	1.01
Shouted	7	21.07	1.32
Landed	6	21.6	1.33
Cried	5	24.66	1.39
Jumped	6	28.43	1.45
Laughed	6	43.5	1.64
Smiled	6	51.55	1.71
Arrived	7	70.48	1.85
Walked	6	104.82	2.02
Died	4	133.27	2.12
Worked	6	167.26	2.22
Called	6	405.44	2.61

Appendix B: Experimental items for testing regulars.

Note that these are all the grammatical versions.

Yesterday they cried.
Yesterday they migrated.
Yesterday he died.
Yesterday she smiled.
Yesterday it rained.
Yesterday it snowed.
Yesterday he jumped.
Yesterday he called.
Yesterday they landed.
Yesterday we laughed.
Yesterday she arrived.
Yesterday we quarreled.
Yesterday I shaved.
Yesterday I worked.
Yesterday he shouted.
Yesterday she moaned.
Yesterday we walked.
Yesterday she frowned.
Yesterday she chuckled.
Yesterday he giggled.