Tanja Angelovska*, Dietmar Roehm and Sabrina Weinmüller

# Uncovering transfer effects of dominance and proficiency in L3 English acquisition using the visual moving window paradigm and grammaticality judgments

https://doi.org/10.1515/applirev-2019-0075

**Abstract:** Using a novel combination of visual moving window paradigm and timed grammaticality judgment task, this study examines how third language (L3) learners (beginners and intermediate) with L2 German and different non-verb-second L1s process violated and non-violated main declarative sentences with fronted adverbials in L3 English. It examines the extent to which so far less-explored predictors (language dominance and proficiency) modulate non-facilitative word order transfer from the L2. Our results from experiment 1 corroborate existing (offline data) results (Angelovska, Tanja. 2017. (When) do L3 English learners transfer form L2 German? Evidence from spoken and written data by L1 Russian speakers. In Tanja Angelovska & Angela Hahn (eds.), *L3 syntactic transfer: Models, new developments and implications* (Bilingual Processing and Acquisition 5), 195–222. Amsterdam, Philadelphia: John Benjamins; Fallah, Nader & Ali Akbar Jabbari. 2018. L3 acquisition of English attributive adjectives dominant language of communication matters for syntactic cross-linguistic influence. *Linguistic Approaches to Bilingualism* 8. 193–216) and are in support of a hybrid transfer suggesting that neither proficiency nor dominance plays a role in transfer selection. Results from experiment 2 reveal that L1-dominance was the determining key factor for accuracy performance for low proficiency L3 subjects but higher L3 proficiency tended to neutralize this strong influence - providing evidence for the Scalpel Model (Slabakova, Roumyana. 2017. The scalpel model of third language acquisition. *International Journal of Bilingualism* 21. 651–665). We explain the contradictory results from the two experiments as a function of task effects.

**Keywords:** third language acquisition, syntactic transfer, visual-moving window, proficiency, dominance

***Corresponding author: Tanja Angelovska**, Department of English and American Studies, University of Salzburg, Salzburg, Austria, E-mail: tanja.angelovska@sbg.ac.at
**Dietmar Roehm,** Department of Linguistics, Centre for Cognitive Neuroscience, University of Salzburg, Salzburg, Austria, E-mail: dietmar.roehm@sbg.ac.at
**Sabrina Weinmüller,** Department of English and American Studies, University of Salzburg, Salzburg, Austria, E-mail: sabrina.weinmueller@sbg.ac.at

# 1 Introduction

In the last decade, there has been an increased interest in studying how the previously learned linguistic systems affect third language acquisition. Several predictors, such as order of acquisition, (psycho-)typology, recency of use, proficiency in the target language and all previously acquired languages, length of residence and exposure to the target language environment (cf. De Angelis 2007) have been considered as constraining the type of transfer in the target (third) language. Several proposals have been brought forth to explain which of the prior languages will take over the leading role in L3 acquisition. While some of these grant a privileged status to one of the existing language systems (e.g., the L1 transfer scenario, Hermas 2014a, 2014b; the L2 Status Factor, Bardel and Falk 2007, 2012; Bardel and Sánchez 2017; Falk and Bardel 2011), others consider it possible that both the L1 and the L2 play a role in L3 acquisition (e.g., the Cumulative Enhancement Model, Flynn et al. 2004; the Typological Primacy Model, Rothman 2010, 2011, 2013, 2015; Rothman and Cabrelli Amaro 2010).

As a response to recent calls for including moment-by-moment measures for transfer in L3 acquisition (cf. Angelovska 2017; González Alonso and Rothman 2017a; Puig-Mayenco et al. 2018; Rothman et al. 2015), we examined the extent to which so far less-explored possible predictors (language dominance and proficiency) moderate the non-facilitative transfer of word order from the L2 while processing L3 sentences by using both an online method (self-paced visual moving window paradigm) and an offline method (timed grammaticality judgment task). The incremental measure in the visual moving window provides a more fine-graded perspective on the issue of transfer. L3 learners of English of various L1 backgrounds and German as L2 read and judged main declarative sentences with fronted adverbials in two conditions (a non-violated condition and a violated condition). In the latter, subjects and verbs were inverted to reflect the German verb-second order. Our study implies additional evidence for an existing gap, addressed by Hopp (2018), who tested whether besides typology and status, the effects of language dominance and proficiency, as different constructs, seem to be additive in L3 acquisition.

The paper is organized as follows: The introduction is followed by a state-of-the-art review on syntactic transfer in L3 acquisition and a review of previous L3 studies with a focus on proficiency and language dominance. Subsequently, we elaborate the target feature (verb second) and the reasons for the inclusion of the visual moving window paradigm and the grammaticality judgment task. Then, we present the methodological aspects of the study and its results. We conclude by discussing the implications and limitations.

## 2 Syntactic transfer in third language acquisition

Undoubtedly, whether one transfers from one, two or more languages, transfer is the most striking difference that distinguishes L3 learners from L2 learners. So far, four scenarios about what is transferred during the initial stage exist (the L1 Factor, Hermas 2010, 2014a, 2014b; the L2 Status Factor: Bardel and Falk 2007, 2012; Bardel and Sánchez 2017; Falk and Bardel 2010, 2011; Falk et al. 2015; the Cumulative Enhancement Model: Berkes and Flynn 2012; Flynn et al. 2004; and the Typological Proximity Model: Cabrelli Amaro et al. 2015; Rothman 2010, 2011, 2013, 2015).

Notwithstanding that a formal model of absolute L1 transfer was never formally proposed, there are a few studies (Hermas 2010, 2014a, 2014b; Jin 2009; Na Ranong and Leung 2009) that account for a full L1 transfer in L3 acquisition. An absolute L1 transfer would suggest that the L1 has some type of privileged status and acts as a sort of filter, blocking access to acquired L2 properties. However, as in these studies, other factors such as structural relationships between the languages have not been considered. Thus, the results could also be explained with one of the other formally proposed models.

According to the cumulative enhancement model, the language acquisition process is non-redundant and cumulative allowing transfer from either of the previously acquired languages. As a result of cognitive economy, the proponents claim that transfer will only occur if it is facilitative.

The L2 status factor hypothesis claims that the L2 is the only possible source of transfer since it acts as a filter, which makes the L1 inaccessible in L3 acquisition. The privileged role of the L2 is based on the differences between L1 and L2 acquisition, for example, age of onset, naturalistic versus formal learning environment, and the degree of metalinguistic knowledge. Bardel and Falk (2012) supported their claims by following Paradis' (2004) model, claiming that L2 grammar is stored in declarative memory (explicit knowledge), while L1 grammar is stored in procedural memory (implicit knowledge) system. Hence, there is a higher cognitive similarity between L2 and L3, which, as a result, makes the L2 more easily accessible in L3 acquisition. The most recent refinements of the L2 status factor (though not attested empirically yet) include such variables as the role of working memory, attention and noticing. Another addition is that the L2 status factor allows L1 transfer under a certain condition; namely, when learners have a high degree of explicit metalinguistic knowledge in the L1 (Bardel and Sánchez 2017; Falk et al. 2015).

The typological primacy model maintains that the L1 and the L2 are possible sources of transfer, hypothesizing that during the initial stages of L3 acquisition, after enough input, the internal parser determines which system should be transferred based on typological proximity and this system is then transferred

holistically rather than gradually, in line with the Full Transfer-Full Access Hypothesis (see Schwartz and Sprouse 1996). Rothman (2013) has introduced an implicational hierarchy of linguistic cues the parser operates on to determine which language to select for transfer: (1) lexicon, (2) phonological cues, (3) functional morphology, and (4) syntactic structure. The body of available empirical evidence supporting this model is large (e.g., Cabrelli Amaro et al. 2015; Giancaspro et al. 2015; Rothman 2010, 2011; Rothman and Cabrelli Amaro 2010; for an overview see Rothman et al. 2018) – all supporting the role of the typologically more proximate language as a source language irrespective of the order of acquisition and regardless of whether transfer is facilitative or non-facilitative.

Recently, other models have been proposed explaining transfer beyond the initial stages in L3 acquisition (e.g., The Linguistic Proximity Model by Westergaard et al. 2017 and the Scalpel Model by Slabakova 2017). Only the latter considers the role of language dominance in transfer effects. However, the question of how we can further explain the interplay of proficiency and language dominance when using online measures remains open.

## 3 Language proficiency and dominance as predictors for transfer in L3 acquisition

Proficiency has been found to influence the source and amount of transfer in L3 acquisition, as attested in studies that rely on language production data (e.g., Angelovska and Hahn 2014; Bardel and Falk 2007; Falk and Bardel 2010). It is generally maintained that transfer is more likely to occur at the lower proficiency level (e.g., Jaensch 2009; Navés et al. 2005; Odlin 1989; Ringbom 1986; Williams and Hammarberg 1998) with facilitative transfer leading to target-like constructions. At a later proficiency level, however, it becomes impossible to distinguish between facilitative transfer and true (not mediated by transfer) acquisition (cf. González Alonso and Rothman 2017a).

One of the few visual moving window studies examined the role of transfer in processing double object constructions in L3 English by a group of Basque/Spanish bilinguals and an L1 Spanish-L2 English group, divided into three L3 proficiency levels (Imaz Agirre and García Mayo 2017). They did not find any differences in accuracy scores and reading times between the intermediate and the advanced group. Some researchers (e.g., Tokowicz et al. 2004) used measures of proficiency to determine language dominance. Later accounts call for caution claiming that one must distinguish between language dominance and proficiency (cf. Birdsong 2006) – as "one can be dominant in a language without being highly proficient in

that language" (Gertken et al. 2014: 209) and "dominance may shift (…) independently of proficiency" (Gertken et al. 2014: 211). Rather than being the sole factor that defines language dominance, proficiency in each language is only one component of language dominance. Which other components make up the construct of language dominance is not an easy task to solve as opinions differ. Birdsong (2006), for example, conceptualizes dominance in terms of processing differences between the L1 and the L2. He states that "an individual's L2 is dominant if, compared to this person's L1, performance on a battery of language tasks is characterized by greater speed, fluency, automaticity, or efficiency (accuracy) in processing" (Birdsong 2006: 47). Heredia (1997) and later Harris et al. (2006), on the other hand, define the dominant language as being the one that is most accessible in day-to-day life, most activated and most frequently used.

Slabakova's Scalpel Model (2017) proposes dominance as a potential factor to influence transfer selection. However, to date, there is only a small number of studies that investigated the role of dominance in relation to syntactic transfer. Rah (2010) conducted a study on transfer effects of the relative clause attachment ambiguity in two groups of German learners of French. Participants in the first group had started to learn English before French and were considered English L2 and French L3 learners, whereas participants of the second group had started to learn French before English and were considered L2 learners of French and L3 learners of English. Language dominance was assessed only via self-ratings. Findings of the study show that English-dominant learners of French transfer the attachment preference from English to French, while French-dominant learners, in contrast, were not influenced by the English preference. Hence, dominance was found to have a greater influence on transfer than the length of exposure to a foreign language. Fallah and Akbar Jabbari (2016) investigated the role of dominance in L3 acquisition of English attributive adjectives by three groups of teenage bilinguals: L1 Mazandarani/L2 Persian speakers who were Mazandarani-dominant (mean age = 13.10), L1 Mazandarani/L2 Persian speakers who were Persian-dominant (mean age = 13.09), and L1 Persian/L2 Mazandarani speakers who were Persian-dominant (mean age = 13). Dominance was determined by assessing which language participants used as the language of communication at home, in social contexts, and at school. Data gathered via a grammaticality judgment task and an element rearrangement task indicate that dominance is the main predictor for syntactic transfer. However, Lloyd-Smith et al. (2018) found that syntactic transfer occurs from both languages and is unrelated to overall dominance in German or proficiency in the heritage language (Italian).

The most recent visual moving window study (Puig-Mayenco et al. 2018), investigating the role of dominance in transfer selection, compared two sets of early Spanish-Catalan-bilinguals with high levels of proficiency in both languages,

who were absolute beginners of L3 English differing in their dominant languages. They examined the syntactic distribution of negative quantifiers in English, a phenomenon that differs in Spanish and Catalan. Their results revealed that dominance was not a significant predictor since in both dominance groups the same language was selected as the source of transfer (Catalan). These results support the typological primacy model, as Catalan is closer to English than Spanish (at least at the phonotactic level). The authors concluded that dominance is not able to trump traditionally assumed variables in determining the source of transfer. Specific studies considering dominance in relation to the verb-second transfer of L2 German in L3 English will be reviewed after a short theoretical introduction to the target feature (verb-second).

## 4 The verb-second phenomenon in L3 acquisition

Probably the most studied phenomenon in Germanic languages within Generative Grammar is the verb-second (Haider 2010; Holmberg 2012; Westergaard 2009). Apart from English, all Germanic languages are verb-second – the phenomenon where the finite verb is required to appear in the second position of a declarative main clause preceded by a single arbitrary constituent. The initial sentence arbitrary constituent is not (necessarily) the subject. In German, the finite verb must always follow the fronted adverbial in a declarative main clause (with usually the subject following the main verb in base position). Regarding transitive clauses, this is only valid when all arguments are full NPs as pronouns jump to the Wackernagel-position and thereby precede subjects (e.g., *Plötzlich sprang ihm die Katze auf den Kopf* vs. *Plötzlich sprang die Katze ihm auf den Kopf*). See Example 1a in comparison to 1b where German requires the verb element to be the second constituent in the main clause independently of the initial constituent:

(1a) Die Katze sprang plötzlich auf den Baum.
S V Adv O

| | | | | |
|---|---|---|---|---|
| (1b) | Plötzlich | sprang | die Katze | auf den Baum. |
| | *Suddenly | jumped | the cat | onto the tree. |
| | Adv | V | S | O |

| | | | | |
|---|---|---|---|---|
| (1c) | Suddenly | the cat | jumped | onto the tree. |
| | *Plötzlich | die Katze | sprang | auf den Baum. |
| | Adv | S | V | O |

It is obvious that in German (see example 1b) the sentence takes the order of adverb-verb-subject (AdvVSO). In English (and in non-Germanic languages without verb-second), in contrast, a main declarative sentence with a fronted constituent, such as an adverb, will not result in a subject-verb inversion, i.e., the English sentence will keep its rigid SVO order, as illustrated in example (1c). English (AdvSVO) allows neither VSO nor XVSO (X representing any single arbitrary constituent).

In L3 acquisition, the number of studies investigating the transfer of verb-second is limited (Angelovska 2017; Bardel and Falk 2007; Hopp 2018; Mykhaylyk et al. 2015; Sağın Şimşek 2006). Sağın Şimşek (2006) was the first to report on the non-facilitative transfer of verb-second in L3 acquisition of English by Turkish-German teenage bilinguals. Results of written production data showed German induced word order features in 233 English sentences out of 465 and out of those 233, 38 were topicalized. All 38 attempts were incorrect, i.e., Turkish-German bilinguals placed the finite verb in the second position in L3 English in main and subordinate clauses, thus applying the German verb-second rule to English sentence structure – due to typological similarity. Similarly, using oral data, Bardel and Falk (2007) looked at the placement of sentence negation in L3 acquisition of Swedish or Dutch (both Germanic languages) by two groups of young adults with different L1s and L2s, whereby one of their previously acquired languages was a verb-second language (Swedish/German/Dutch), while the other one was not (English/Hungarian/Italian/Albanian). They concluded that placement of negation was more easily transferred from the L2 than from the L1, assigning the L2 a privileged role as a source of transfer at the L3 initial state.

Angelovska (2017) and Hopp (2018) investigated the transfer of the verb-second. Angelovska (2017) examined the occurrence of a non-facilitative transfer of the verb-second by adult L3 learners of English at different proficiency levels (beginners and intermediates) with L1 Russian and L2 German. Spoken and written data was collected utilizing an oral storytelling task and a written narrative task. Both tasks targeted the elicitation of main declarative sentences with fronted temporal adverbials to examine the occurrence of a negative transfer of verb-second from L2 German into L3 English. Additionally, dominance was taken into consideration as a potentially confounding factor for transfer. It was determined by using a language-use-based dominance scoring procedure based on a questionnaire that provided information about the participants' use of the languages under consideration. For data analysis, Obligatory Occasion Analysis was used, i.e., a percentage of accurate use was calculated. The results showed that L2--dominant beginners had a low score in both modalities and intermediate L2--dominant subjects displayed low scores in writing, implying that they produced verb-second sentences. These results confirmed the hypothesis that when negative

interlanguage transfer of verb-second from L2 German in L3 acquisition of English occurs, this transfer is dependent on dominance. Hopp (2018) included data from bilingual children comparing two groups of Turkish-German bilingual and German monolingual children learning English. Hopp (2018) used a sentence repetition task and a picture story-retelling task and focused on verb-second and adverb order, as well as on verb-complement order and article and subject realization. He determined dominance based on vocabulary scores in German and Turkish only, without any measures for proficiencies in the three languages – something he reports as a limitation himself. The data displayed consistent similarities between the two groups tested, namely that both transferred from their dominant L2 – a result reported as compatible with the typological primacy model. Hopp (2018) raised the question whether "transfer from German reflects effects of typology or dominance or whether their effects are additive" (Hopp 2018: 14; see also Slabakova 2017; Westergaard et al. 2017). To account for possible additive effects, more subtle data is needed.

## 5 Previous studies with the visual moving window paradigm and timed grammaticality judgments

With the visual moving window technique, processing difficulties, arising due to ambiguities, anomalies or distance dependencies, are examined. In our study, the targeted phenomena are anomalies. In this technique, the time that elapses between two successive presses of the trigger is recorded. The basic assumption is that the time a participant needs to read a segment reflects the cognitive workload of processing that segment, thus indicating sensitivity to the linguistic phenomenon tested (Marsden et al. 2017). Longer reading times are, therefore, interpreted as processing difficulties, while faster reading times are a sign of facilitation (Jegerski 2014; McDonough and Trofimovich 2012). Processing difficulties can primarily be seen in increased reading times at or after the violation and the sentence-final word is usually included for the detection of spillover and sentence wrap-up effects (see Jegerski 2014). Inconsistency regarding the regions was found in L2 studies. For example, De Vincenzi et al. (2003) reported significantly longer reading times for the critical word and the word following it, but no differences in the words preceding the critical one. Similarly, Tokowicz and Warren (2010) and Jiang et al. (2011) reported increased reading times at the point of violation (i.e., critical word) and the word(s) following it, whereas Roberts and Liszka (2013) did not find longer reading times at the critical word itself but at the words following it. Ditman et al. (2007) reported increased reading times only at

the point of violation. Some have considered grammaticality judgments as assessing implicit knowledge if the task is timed because participants would judge the sentences intuitively and would not have time to access their explicit knowledge. Several studies (Bowles 2011; Ellis 2005; Godfroid et al. 2015) have investigated the use of grammaticality judgments in relation to implicit or explicit knowledge. Ellis (2005) designed a battery of tests, examining whether they measured implicit or explicit knowledge. He reported that L2 learners of English performed worse when they did a timed grammaticality judgment task than an untimed one since, without the time pressure, they had time to supplement their implicit with explicit knowledge – concluding that only a timed task ensures that implicit knowledge is activated. Bowles (2011) replicated this study with L2 learners of Spanish and included a group of Spanish heritage speakers. Results showed that heritage learners with very little explicit knowledge performed better on the timed task. Godfroid et al. (2015) also conducted two grammaticality judgment tasks, timed and untimed, to gather complementary data. Their findings contributed further evidence for the hypothesis that grammaticality judgments with and without time constraints correspond to implicit and explicit knowledge, respectively. Although these studies included L2 participants, we can assume that the same methodology will prove beneficial for accessing implicit knowledge in L3 acquisition, likewise. Despite numerous visual moving window studies in L2 research (for a comprehensive review see Marsden et al. 2017) investigating the difference between native and non-native processing, this method has barely been used in L3 transfer studies (Westergaard et al. 2017 is an exception although they included a timed Powerpoint Presentation and no details were given about the timing).

# 6 Hypotheses

To date, neither the visual moving window technique nor timed grammaticality judgments have been employed to examine syntactic violations in L3 acquisition. Thus, in line with our review, the following hypotheses regarding reading times, reaction times, and accuracy data were generated:

**H1:** English native speakers will have shorter reading times for each region of interest and shorter reaction times and higher accuracy rates than L3 learners of English when judging sentences in both conditions (violated and non-violated).

**H2:** When reading ungrammatical sentences, L3 beginners of English will have shorter reading times. When judging ungrammatical sentences, they will show

longer reaction times and lower accuracy rates than intermediate learners of English.

**H3:** When reading and judging ungrammatical sentences, L1-dominant and L2-dominant learners will show differences in reading times, reaction times, and accuracy rates.

For H1 and H2 predictions were derived from the typological primacy model, whereas H3 refers to Slabakova (2017). Only hypothesis 1 refers to possible differences between NS and L3 learners whereas hypotheses 2 and 3 focus on the L3 learners only and test within-group differences.

# 7 The present study

## 7.1 Participants

Thirty subjects were recruited through printed and electronic advertisements on notice boards at two universities in western Austria and southern Germany, where they resided at the time of the testing. The English native speakers, who served as a control group ($n = 15$, eight female), were students who were on a temporary stay in Austria, with a mean age = 20.7 (SD = 3.9) with absolutely no knowledge of German and a stay in Austria of no longer than two weeks – which was the prerequisite for being included in the native speakers' group. The L3 group was comprised of advanced speakers of L2 German who have been living in a German-speaking country (for 9.4 years on average) at the time of testing, having acquired English as a second foreign language (chronologically L3), subsequently after German and had different non-verb-second L1 backgrounds (L1x). They were L1x-L2 German successive bilinguals with L3 English ($n = 15$, 12 females, mean age = 27.87, SD = 7.86, minimum age = 20 and maximum age = 47). For example, their L1s were Arabic, Bosnian, Croatian, Georgian, Kurdish, Mongolian ($n = 3$), Polish, Slovakian, Slovenian, Spanish, Turkish, and Vietnamese ($n = 2$). At the time of the data elicitation, they were all students at a German-speaking State University having passed the official German-language entrance exam that requires them to have reached the C1/C2 advanced level in German (as determined by the *Common European Framework of References for Languages*). They all started learning German and English at different ages, with a mean age of onset 11.9 and 12.1 years, respectively. None of the participants reported any reading problems. They participated voluntarily, and all signed an informed written research consent form prior to the experiments.

## 7.2 Procedure

Participants were tested individually on a Windows 10 IntelR Core i-5 CPU computer with a 13-inch HD monitor (60 Hz refresh rate). The experiment was designed and run with PsychoPy software (Peirce 2007, www.psychopy.org) and reading and reaction times were measured via a keyboard press. Participants were seated in a quiet environment at approximately 60 cm from the computer screen. There was a short break between the two experimental tasks, which took approximately 20 min to complete.

Prior to the experiment, under the observation of one of the authors, participants performed the vocabulary and grammar parts of the *Oxford Placement Test* to assess their proficiency in English. The most frequently used technique for measuring proficiency are objective standardized placement tests. The proficiency test allowed an automatic assigning of the proficiency level (A1 to C2), including the total number of scores achieved for both sections per subject (total possible for both grammar and vocabulary was 40 each). Afterwards, they completed a language background questionnaire to determine their dominance and to collect biographical data. The German version of the language background questionnaire was adapted from the *Bilingual Language Profile Tool* (Birdsong et al. 2012). The questionnaire was made available through the *Google survey* tool. Based on participants' answers, dominance was calculated using the scoring procedure recommended by the *Bilingual Language Profile*. First, global language scores for each prior language (L1 and German) were calculated separately and then one language total (German) was subtracted from the other (L1) to obtain the dominance score. Scores smaller than –1 indicated an L2-dominance, while scores above one indicate L1-dominance. A score of zero would reflect balanced bilingualism. The dominance score was taken as a continuous variable.

## 7.3 Experiments and material

### 7.3.1 Experiment 1: The visual moving window technique

Before the experiment started, the L3 learners' group was provided with a word list to eliminate the possibility that longer reading times for specific words were due to processing difficulties or lack of knowledge of the lexical items themselves (Juffs and Rodriguez 2015: 35). We ensured that non-native English participants understood all words by providing German translations. Right before the start of the task, instructions for the task were provided in German for L3 learners and English for the native speakers' group. They were encouraged to read the sentences silently for

meaning, at their own pace. The instructions emphasized that participants had to press the spacebar to advance to the next word until they reach the end of a sentence. To make participants familiar with the technique, the instructions were followed by a practice trial, which consisted of eight filler sentences, unrelated to the research objective. At the beginning of each sentence, a fixation point indicated where the sentence would start. Each sentence was presented word by word, using the moving-window technique (Just et al. 1982). The stimuli were presented using a non-cumulative (Marinis 2003) linear display allowing "the experimenter to identify the specific loci of processing difficulty" (Kroll et al. 2008: 125).

Eighteen short main declarative sentences with fronted adverbials were created as targets. They were balanced according to length and correctness and were presented in a randomized order. Each sentence existed in two conditions (violated and non-violated), with the violation occurring at the region after the fronted adverbial, as illustrated in example (2b). The sentences in the two conditions were lexically matched (identical words, but different word order) to ensure maximal comparability (Keating and Jegerski 2015: 5). Each of the 18 sentences in both conditions consisted of six words, starting with a fronted adverbial, followed by either the personal pronoun "he" (correct/non-violated condition) or a regular past simple tense verb (incorrect/violated condition).

(2) a. Yesterday he watched an interesting movie.
b. *Yesterday watched he an interesting movie.

Additionally, following Jegerski (2014), 36 filler sentences, not related to the target feature, entailing different grammatical violations of no specific target, were included.

Two counterbalanced presentation lists were created to reduce the risk of participants becoming aware of the targets and to counteract individual differences in reading speed (cf. Jegerski 2014). Thus, the two presentation lists each contained nine target stimuli in the grammatical condition and nine in the ungrammatical, and 18 grammatical and 18 ungrammatical fillers, which adds up to 54 sentences the participants had to read in the visual moving window experiment. The number of regions of interest per presentation list was 108. In line with previous studies from L2 research (De Vincenzi et al. 2003; Jiang et al. 2011; Tokowicz and Warren 2010), the following regions of interest will be analysed: the first region of interest (the fronted adverbial), the second region of interest (the critical word: S/V), the third region of interest (the word following it: V/S) fourth region of interest (pre-pre-final word), the fifth region of interest (pre-final word), and sixth region of interest (the sentence-final word). Note that respectively the second and third region of interest (Verb or Subject) will differ in the two conditions (violated or non-violated).

The program recorded reading times in milliseconds for every keypress and once all sentences were read, the experiment stopped automatically. A short break was included between the visual moving window and the grammaticality judgment experiment.

### 7.3.2 Experiment 2: Grammaticality judgments

The aspect that participants judge will be referred to as "grammatical acceptability" (Chomsky 1965: 11). This experiment was administered using the same computer and software as in experiment 1. Before the experiment started, detailed instructions were presented (same language choice as in exp. 1). The experimental task was first piloted with five native speakers and their average response time (=2.7 s) was taken as a measurement to which 70% was added to calculate the time limit for the L3 group, resulting in 5 s (cf. Bowles 2011; Ellis 2005; Godfroid et al. 2015). At first, 50% were added because the L3 group were beginners or intermediate learners. After piloting the task with beginners, the time given was still not enough, since the pilot participants missed many trials; hence, another 20% of the time was added to the mean reaction time (henceforth, RT) of the native speakers, finally allowing 5 s for judging the sentences for the L3 group. The pilot subjects were discarded from the pool. The instructions were followed by a practice trial of eight fillers (not related to the research objective). Participants were asked to judge the sentences according to their grammatical acceptability, using a binary response scale (Yes/No), because the stimuli sentences present a clear and strong grammaticality contrast and a rating scale with more than two options would only be appropriate for more subtle contrasts (cf. Ionin 2012: 42). Sentences were displayed one at a time. To judge a sentence, participants had to press 'y' or 'n' on the keyboard (tag-labelled), respectively. The program automatically recorded the responses and RT. If no judgment was made within the 5 s, the program automatically moved on to the next sentence and no response was recorded.

Since the order of presentation of the stimuli can affect participants' judgments, the sentences were randomized using the randomization function of the software. This means that each participant read and judged the sentences in a different order, and it was avoided that judgments were influenced by factors such as nervousness at the beginning of the experiment, fatigue towards the end, the influence of surrounding test items, and practice effects (Schütze 2016:180).

Following Jegerski (2014: 34), the grammaticality judgment task was constructed using stimuli that were different from the visual moving window stimuli, in contrast to Roberts and Liszka (2013), who also employed the two tasks independently, but with identical stimuli. Our timed grammaticality judgment task consisted of 36 (18 in each condition) target sentences – matched in length and

lexical and syntactic complexity. The sentences were divided equally into two conditions (grammatical/correct/non-violated and ungrammatical/incorrect/violated) not to mislead participants into expecting more sentences in one or the other condition. Each of the grammatical sentences began with an adverbial, followed by a subject and third-person singular present simple or a past simple verb. In the ungrammatical sentences, the position of the subject and the verb was reversed. All target sentences in a correct and violated condition (not including the filler sentences) consisted of six words. Following Ionin (2012), Schütze (2016) and Tremblay (2005), an equal number of filler sentences (7–9 words), testing different phenomena, were included in each condition ($n$ = 36 in total).

# 8 Data analysis

For experiment 1, reading time data were trimmed to minimize the effects of outliers. Following Jegerski's (2014) recommendation, a combination of the absolute cut-off and standard deviation method was used for the data trimming process. First, absolute cut-offs were used to identify the most dramatic outliers. The dataset was screened for reading times below 250 ms and these outliers were removed from the data, resulting in an elimination of 3.95% of the raw reading times (native speakers group: 2.99%; L3 group: 0.9%). Next, outliers that were more than three standard deviations above the single subject/single item mean were removed, which affected 2.86% of all remaining data (native speakers group: 0.64%; L3 group: 2.22%).

For experiment 1 and 2, we performed linear and general linear mixed model (henceforth, LMM/GLMM) analyses using the *R* (R Core Team, 2020, Version 4.0.2) package *lme4* (Bates et al. 2015, Version lme4_1.1-23). Using the R package *car* (Fox and Weisberg 2019, Version carData_3.0-9), categorical variables were encoded with sum contrasts, such that the estimate for a given fixed effects level represents the difference between this level and the grand mean (for a detailed discussion of contrast coding, see Schad et al. 2020). All models included random intercepts for both subjects and items. While the use of maximal random effects structures has been encouraged for psycholinguistic research by some authors (Barr et al. 2013), we chose to use random intercepts only for the present study because we focus on population-level effects rather than inter-individual differences.

We calculated the full model summary provided by the *lme4* package (Bates et al. 2015, Version 1.1-23), results of Type-II Wald $\chi^2$ tests provided by the *car* package (Fox and Weisberg 2019, Version 3.0-9) and estimated marginal means from the *emmeans* package (Lenth et al. 2020, Version 1.4.8) derived from our model. For the visualization of effects of interest, we used the following R

packages: *car* (Fox and Weisberg 2019, Version 3.0-9), *ggplot2* (Wickham 2020, Version 3.3.2) and *ggeffects* (Lüdecke et al., 2020, Version 0.15.1). The plots also serve to resolve interactions.

In experiment 1 mean reading times were used as a dependent variable for the LMM analyses. As fixed factors, we included the factors REGIONS OF INTEREST (with the levels ADV = fronted adverbial, S = subject/verb, V = verb/subject, PRE-PRE = pre-pre-final word, PRE = pre-final word, and FINAL = final word), CORRECTNESS (correct, incorrect), and the between-subject factor GROUP (native speakers, L3).

To investigate whether there is an interaction between proficiency, and dominance concerning transfer, a separate LMM analysis was computed including only the reading times of the L3 group. In this analysis, we included the continuous factors PROFICIENCY and DOMINANCE as predictors in addition to the categorical factors REGIONS OF INTEREST (ADV = fronted adverbial, S = subject/verb, V = verb/subject, PRE-PRE = pre-pre-final word, PRE = pre-final word, and FINAL = final word), and CORRECTNESS (correct, incorrect).

In experiment 2, mean accuracy scores with log-transformed reaction time as a predictor were subjected to GLMM analyses including the fixed factors CORRECTNESS (correct, incorrect) and the between-subject factor GROUP (native speakers, L3). Again, separate GLMM analyses were computed, including only the L3 data. In addition to the fixed factor CORRECTNESS, offline values of language proficiency and language dominance were included as continuous predictors (PROFICIENCY and DOMINANCE ).

# 9 Results

## 9.1 Experiment 1

The best fitting model for the data includes position, correctness and group as fixed factors as well as by-subject and by-item random intercepts but no random slopes (by-subject, by-item) [lmer (formula mean ~ 1 + pos * corr * group + (1 | subj) + (1 | item), data: data1, REML = F)].

Statistical analysis of the reading time data revealed main effects of position [type II Wald test: $\chi^2(5) = 114.01$, $p < 0.001$], correctness [$\chi^2(1) = 11.27$, $p < 0.001$] and group [$\chi^2(1) = 6.45$, $p < 0.05$] , as well as an interaction between position and group [$\chi^2(5) = 51.74$, $p < 0.001$]. Model estimates are visualized in Figure 1 using estimated marginal means. This also serves to resolve the interaction. The shaded areas in this and the following figures represent 83% confidence intervals, the non-overlap of which corresponds to significance at the 5% level.

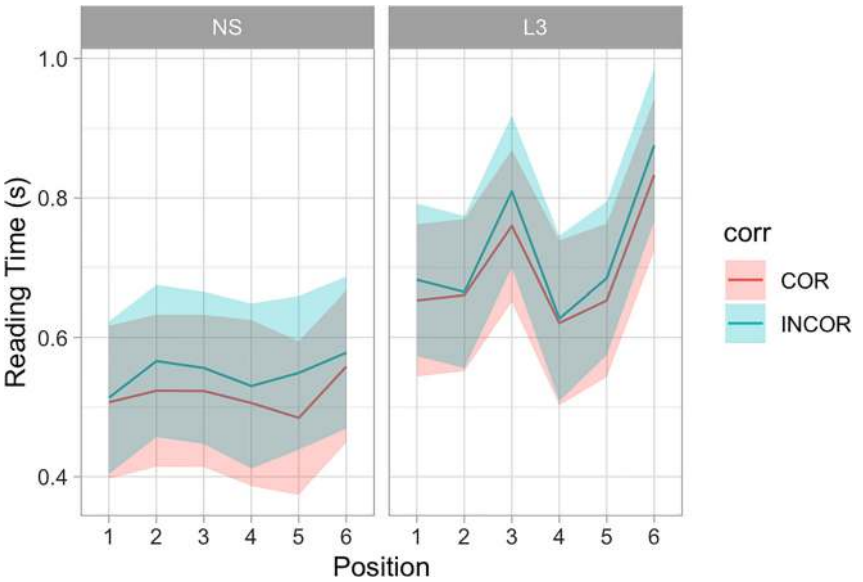


**Figure 1:** Visualization of the interaction between correctness, position, and group. Shaded areas indicate 83% confidence intervals (the non-overlap of which corresponds to significance at the 5% level) computed by the Effect() function from the *effects* package (Fox and Weisberg 2019). Abbrev. Position: 1 = ADV, position 2 = critical word (S/V), position 3 = the word following the critical word (V/S), position 4 = pre-pre final word, position 5 = pre-final word and position 6 = sentence-final word.

For the analysis of the L3 group iterative model fits revealed that the base model (including only the fixed factors) showed the best fit to the data [lmer (formula = mean ~ 1 + pos * corr + (1 | subj) + (1 | item), data: L3, REML = F]. Inclusion of the predictors DOMINANCE or/and PROFICIENCY to the base model and random intercepts for participants and items showed no improvements compared to the base model (comparison base model to Dominance model: $\chi^2 = 0.17$, $p = 0.68$; comparison base model to Proficiency model: $\chi^2 = 0.04$, $p = 0.99$). Hence reading time was not affected by those factors. Type II Wald tests revealed a main effect of position [$\chi^2(5) = 109.56$, $p < 0.001$], and a marginal effect of correctness [$\chi^2(1) = 3.68$, $p = 0.055$].

## 9.2 Experiment 2

The mean accuracy rate for the NS group revealed for the correct condition 94.1% (sd = 23.65%), for the violated condition 87.4% (33.24%) and for the L3 group

86.0% (sd = 34.77%) in the correct condition and 60.4% (sd = 49.00%) for the incorrect condition. Generalized linear mixed models were performed on accuracy data. Iterative model fits revealed that the best fitting model for the data includes log-transformed RT as predictor and by-subject and by-item random intercepts as well as by-subject random slopes for correctness, and by-item random slopes for group [glmer(corr ~ 1 + log(rt + 1)*con*group + (1 + con | subj) + (1 + group | item)].

The type II Wald test revealed main effects of RT [$\chi^2(1) = 3.96$, $p < 0.05$], correctness [$\chi^2(1) = 6.06$, $p < 0.05$] and group [$\chi^2(1) = 9.98$, $p < 0.01$], as well as an interaction between RT, correctness and group [$\chi^2(1) = 6.84$, $p < 0.01$].

For the analysis of the accuracy of the L3 group the base model with the fixed factor correctness including log-transformed RT and continuous dominance and proficiency values as a predictor, by-subject and by-item random intercepts as well as by-subject random slopes for correctness showed the best fit to the data [glmer(corr ~ 1 + con*log(rt + 1) * scale(dom)*scale(prof) + (1 + con | subj) + (1 | item)].

(base vs. dominance model: $\chi^2 = 24.17$, $p < 0.001$; base vs. proficiency model: $\chi^2 = 24.61$, $p < 0.001$; base vs. dom*proficiency model: $\chi^2 = 53.73$, $p < 0.001$; proficiency vs. dom*proficiency model: $\chi^2 = 29.12$, $p < 0.001$; dom vs. dom*proficiency model: $\chi^2 = 29.60$, $p < 0.001$; dom*proficiency vs. inclusion of by-subject random slopes for correctness: $\chi_2 = 23.02$, $p < 0.001$).

Type II Wald tests revealed a main effect of correctness [$\chi^2(1) = 5.28$, $p < 0.05$], as well as an interaction between correctness and dominance [$\chi^2(1) = 4.41$, $p < 0.05$], and correctness, RT, dominance and proficiency [$\chi^2(1) = 10.18$, $p < 0.01$]. Figure 3 to 5 serve to resolve the interaction.

# 10 Discussion

To examine whether and to what extent transfer from the L2 in L3 sentence reading comprehension is subject to effects of language proficiency and dominance, we had L1x–L2 German successive bilinguals and an NS English control group read verb-second and non-verb-second sentences in L3/L1 English preceded by a fronted adverbial using the self-paced non-cumulative visual moving window technique and a timed grammaticality judgment task.

## 10.1 Experiment 1

In hypothesis 1, we predicted that the English NS group would process all sentences in both conditions (correct and incorrect) faster (with shorter reading times for each region of interest). We hypothesized that the NS group would show shorter

reading times in the violated condition because the possibility that German transfer could be obtained is non-existent (native speaker competence in German is not available). As hypothesized, the NS group showed overall faster reading times compared to the L3 group. However, both groups showed faster reading times for correct versus incorrect sentences. Regarding reading times of specific sentence constituents, we found clear group differences. Whereas the NS group showed an increase for the incorrect condition starting at the critical verb that continued over the remaining sentence, the L3 group showed a more specific pattern. While reading times differences (incorrect vs. correct condition) were similar to the NS group, the L3 group showed significantly increased reading times for both conditions at the third sentence position (i.e., the word following the critical word [V/S]) and the sentence's final word. Our results confirm results from previous studies, which demonstrated that increased reading times at the point of violation reflect processing difficulties (De Vincenzi et al. 2003; Ditman et al. 2007; Tokowicz and Warren 2010) and longer reading times for the following regions and the final word are caused by spillover and sentence wrap-up effects (Jegerski 2014: 44). The "unspecific" (i.e., non-condition sensitive) reading time increase for the L3 group might be a result of hybrid transfer from the L1 and the L2 as they were most probably engaged in resolving the correct word order. Such a conflict situation is due to L2 transfer which is caused by the two unstable language systems (English and German) and also resulting in spillover effects. To find out whether these effects were moderated by proficiency and dominance, we did a separate analysis for the L3 group where we controlled for these two factors. Our results showed that proficiency and dominance did not affect reading times.

## 10.2 Experiment 2

Regarding accuracy and RT when judging sentences in both conditions in L3 English, we predicted that NS would judge all sentences faster (shorter RT) and obtain higher accuracy scores. Our results show that there is a complex interaction of different factors (as indicated by the significant interaction between RT, correctness and group; cf. Figure 2).

The two groups showed an inverse pattern. For the NS group, a dependence (negative correlation) between accuracy and RT was only found for correct sentences (the less accurate judgments in the violated condition, the longer RT). For the L3 group, the opposite pattern was evident (i.e., a negative correlation only for incorrect sentences but not for correct ones.) Note that the observed negative correlations cannot be the result of a typical speed-accuracy trade-off effect because otherwise, one would expect exactly the opposite pattern (the less

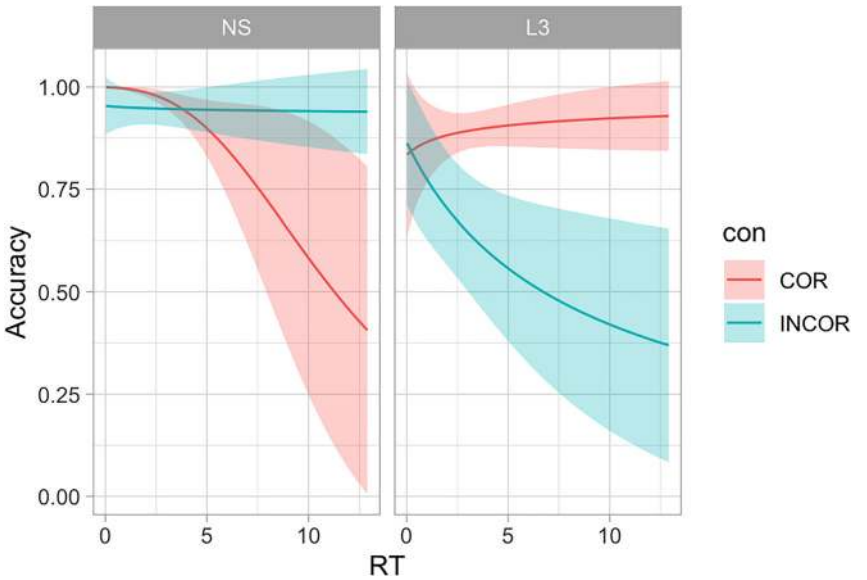


**Figure 2:** Visualization of the interaction between correctness and group in dependence of accuracy and RT. Shaded areas indicate 83% confidence intervals computed by the *Effect*() function from the effects package (Fox and Weisberg 2019).

accurate judgments in the violated condition, the *faster* RT). Hence, this indicates an increased processing difficulty that led to slower decisions and more incorrect judgments for both groups. The processing difficulty with sentences in the violated condition for the L3 group is clearly due to the typological L2 transfer – accounting for the typology primacy model. However, the rather surprising processing difficulty with correct sentences, which was found for the NS group with no knowledge of German might be explained with the type of stimuli and the nature of the judgments themselves. The random effects structure of the model (random by-subject and by-item slopes) reveals subject- and item-specific variation in the NS group. Because fronted temporal adverb phrases in non-emphatic declarative sentences are marked, it could be that some NS subjects considered some of the sentences as not well-formed (without an appropriate context).

Concerning the L3 group, we made predictions about proficiency and dominance assuming that judgments in the violated condition by L3 beginners will reveal longer RT and lower accuracy rates than intermediate learners of English (hypothesis 2) and L1-dominant and L2-dominant learners will show differences in RT and accuracy rates (hypothesis 3). Regarding the influence of proficiency and dominance, the analysis of the L3 group data revealed a complex pattern.

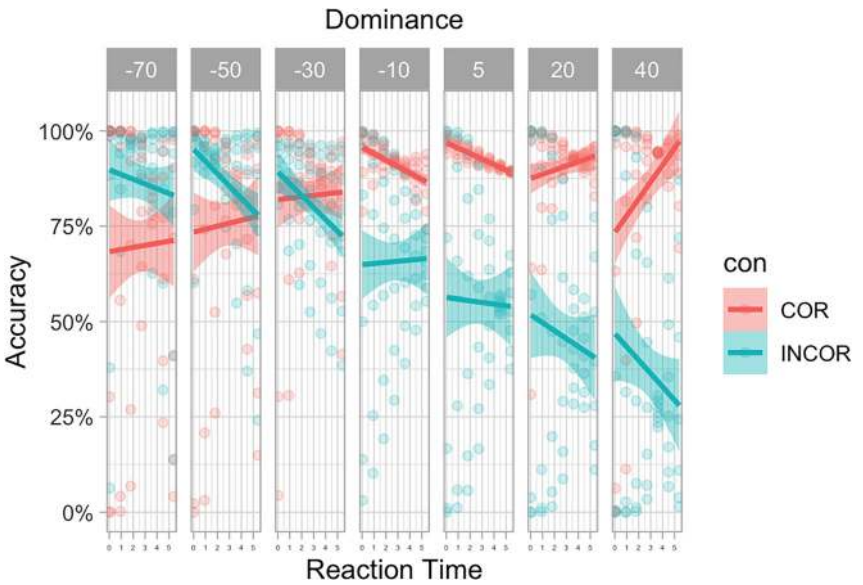


**Figure 3:** Estimated distribution of accuracy values (*y*-axis) in dependence of RT (*x*-axis), dominance values (columns) and correctness (colour). Regression lines with 83% confidence intervals are computed by the *stat_smooth*() function from the ggplot2 package (Wickham 2020, Version 3.3.2). Negative dominance values indicate L2-dominant L3 subjects, positive values indicate L1-dominant L3 subjects.

For correct sentences, there was no clear effect of proficiency although beginners showed more variation in accuracy then intermediate "learners" (cf. Figure 4). However, this variance could be the result of a speed-accuracy effect, showing greater uncertainty in judging correctly and time pressure (interestingly this SAT effect is gone for intermediate and advanced learners). The picture for incorrect sentences is more straightforward. Low proficient subjects had substantial problems in judging incorrect sentences as ungrammatical with increasing proficiency becoming more and more accurate, corroborating earlier findings by Jaensch (2009).

The results for dominance showed little effect on RT but a clear interaction with accuracy that was evident in such a way that for incorrect sentences a decreasing accuracy for the more L1-dominant subjects was found (cf. Figure 3). The opposite was the case for grammatical sentences. L2-dominant subjects judged correct sentences unacceptable whereas increasing L1-dominance paralleled improved performance. This rather surprising result could be interpreted with the weaker language (in this case German) being less dominant and

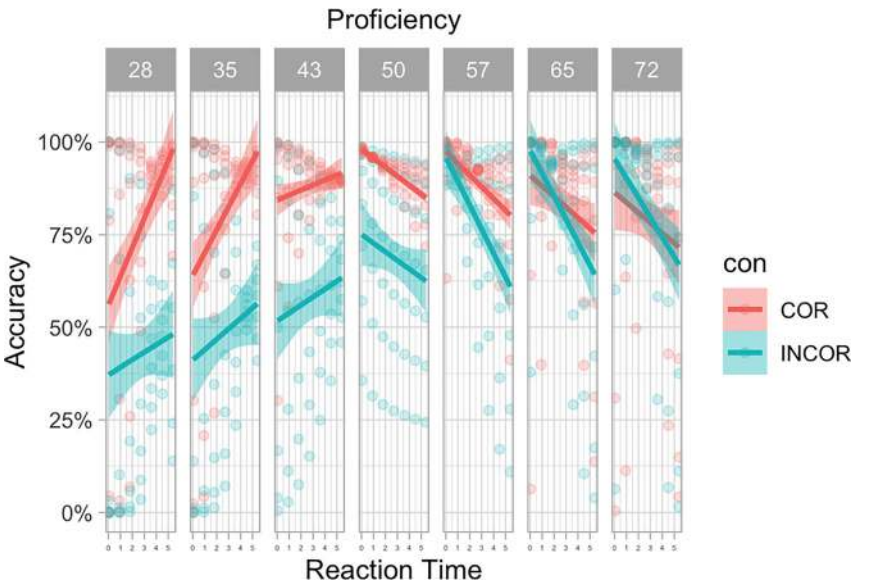


**Figure 4:** Estimated distribution of accuracy values (*y*-axis) in dependence of reaction time (*x*-axis), proficiency values (columns) and correctness (colour). Regression lines with 83% confidence intervals are computed by the *stat_smooth*() function from the ggplot2 package (Wickham 2020, Version 3.3.2).

automatically more likely to be transferred to another less weak and less proficient L3 (in this case English) than a more dominant and stable L1. In other words, the more dominant L1x was more successfully inhibited when the L2-like structure from the less dominant L2 was encountered in the L3. Such inhibition seems to be handled more easily when subjects are required to provide judgments (i.e., the possibility to control is higher than when "just" reading).

Finally, the interaction of dominance and proficiency concerning accuracy revealed interesting dynamics (cf. Figure 5). It shows that dominance was the determining key factor for accuracy performance for low proficiency L3 subjects. Whereas low proficiency L2-dominant subjects showed poor accuracy performance for both correct and incorrect conditions, L1-dominant subjects scored on the ceiling for correct sentences and showed very poor accuracy for incorrect sentences. However, Figure 5 reveals that the higher L3 proficiency was, the more this strong influence of dominance was neutralized and subjects showed higher accuracy in both conditions. Nevertheless, an influence of dominance was still visible even for intermediate subjects and for incorrect sentences (but not for correct sentences). For L2-dominant subjects, the influence of proficiency was

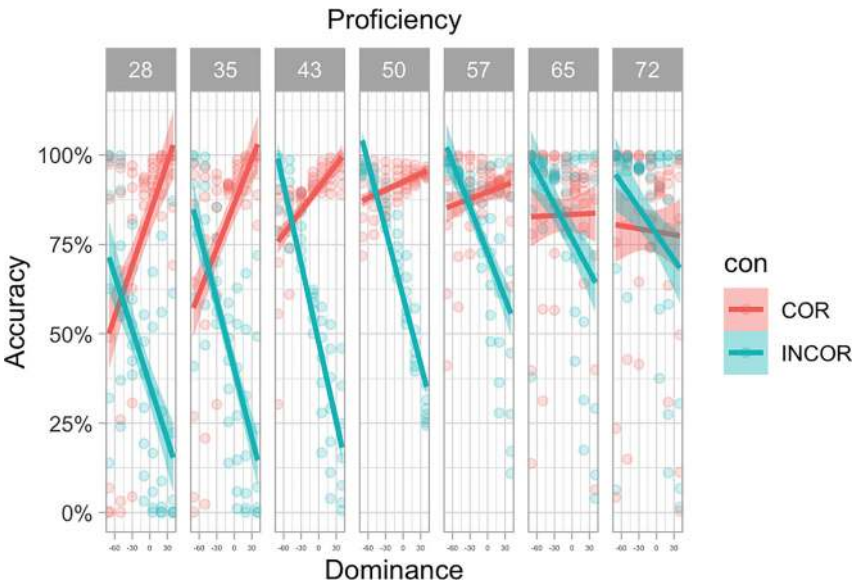


**Figure 5:** mated distribution of accuracy values (*y*-axis) in dependence of dominance (*x*-axis), proficiency values (columns) and correctness (colour). Regression lines with 83% confidence intervals are computed by the *stat_smooth*() function from the ggplot2 package (Wickham 2020, Version 3.3.2).

much stronger. Accuracy performance increased for both conditions in correlation to an increase in proficiency, such that even a slightly above average proficiency total score (=50) resulted in high judgment accuracy for both conditions. Interestingly, the performance advantage for L2-dominant subjects for incorrect vs. correct sentences persisted independently of their proficiency score.

Contra to Hopp (2018) who found that transfer happened from the dominant L2, results from previous studies (Puig-Mayenco et al. 2018) showed that dominance was an indeterministic factor in transfer selection. However, our results showed that dominance is a key factor although in dependence (interaction) of the proficiency level. Whereas low proficient subjects showed a strong influence of dominance, intermediate proficient subjects showed little or no influence. Crucially, in contrast to Hopp (2018), L2 transfer in L3 English was obtained mainly by the L1-dominant subjects. The systematic effect of dominance in our offline measure is ground to claim that dominance is a decisive predictor (cf. Angelovska and Hahn 2012; Angelovska 2017; Fallah and Akbar Jabbari 2016; Rah 2010) accounting for the possibility that dominance might moderate the transfer effects from the competing L1 and L2 – evidence for the Scalpel Model (Slabakova 2017) and in line with recent findings by Puig-Mayenco et al. (2020).

In sum, hypotheses 2 and 3 could be partly confirmed. Whereas for reading times there was not any moderating influence of proficiency and/or dominance, results from accuracy judgment and RTs are in line with H2 and H3.

## 11 Limitations and conclusion

The present study offers new evidence for syntactic transfer in L3 acquisition by examining the effects of dominance and proficiency, as moderating factors on the syntactic non-facilitative transfer of verb-second from L2 German in L3 English.

Our results from experiment 1 cannot be explained solely by the typology primacy model (e.g., Rothman 2015) as the L3 group was struggling to decide about the correct word order (i.e., an indication that the L2 transfer exclusively did not obtain but hybrid transfer [Angelovska and Hahn 2012; Angelovska 2017; Fallah and Akbar Jabbari 2018] did). Because we did not include any measures for metalinguistic knowledge (which is not a construct measured through the visual window paradigm), the L2 status factor cannot be related to our findings. Importantly, results from the second experiment revealed a complex proficiency-dominance pattern indicating that proficiency cannot be looked at independently from dominance.

The contradictory results from the two experiments can be explained with the two different methods (online/offline) measuring two different constructs. We are aware that the time limit for the grammaticality task may have been too long, thus not measuring implicit knowledge. Thus, by activating their explicit knowledge, the L3 group could have applied the correct rules and achieved higher accuracy scores. It remains questionable whether they would have scored differently if given less time. Future studies should test this.

Another limitation of our study refers to the recruitment of the L3 learners. According to mirror-image-design, where L1 and L2 alternate, while the L3 is held constant (e.g., González Alonso and Rothman 2017b), the present study should have included a matched group of L3 learners of English, with L1 German and a non-verb-second L2, residing in the countries where their different L2s are official languages – practically an impossible endeavor. However, Rothman and colleagues showed that variables such as order of acquisition could not be considered as the most explanatory variables at all (cf. Rothman et al. 2018). Ideally, we should have included a third group of advanced L3 subjects (with an equal proportion of L1x-dominant and L2-dominant) to account for how the transfer from the L2 in L3 is mediated by proficiency and dominance across the complete learning trajectory (cf. Cunnings 2017). We are aware that the L1 and L2 of the L3 group might not have the target-like representations considering that they are mostly heritage speakers.

Ideally, the property under investigation should have been tested in their L1 and L2. But since they all have different non-verb-second L1s, it would have been necessary to have one group of native speakers to help with the design of the stimuli in all these different languages and a different group of native speakers to serve as controls. However, the practical reality of finding significant numbers of participants meeting all the needed criteria is different from idealized theoretical propositions.

**Ethics Statement:** All participants gave written informed consent in accordance with the Declaration of Helsinki, and the American Psychological Association's Ethical Principles of Psychologists and Code of Conduct. The protocol was approved by the Institutional Review Board of our University (Code: EK-GZ 07/2018).

## References

Angelovska, Tanja & Angela Hahn. 2012. Written L3 (English): Transfer phenomena of L2 (German) lexical and syntactical properties. In Danuta Gabrys-Barker (ed.), *Crosslinguistic influences in multilingual language acquisition*, 23–40. Heidelberg: Springer.

Angelovska, Tanja & Angela Hahn. 2014. Raising language awareness for learning and teaching L3 grammar. In Alessandro G. Benati, Cécile Laval & María Arche (eds.), *The grammar dimension in instructed second language learning*. In Advances in instructed second language acquisition research book series, 185–207, Alessandro G. Benati (ed.). ISBN 9781474243360. London: Bloomsbury Academic.

Angelovska, Tanja. 2017. (When) do L3 English learners transfer form L2 German? Evidence from spoken and written data by L1 Russian speakers. In Tanja Angelovska & Angela Hahn (eds.), *L3 syntactic transfer: Models, new developments and implications* (Bilingual Processing and Acquisition 5), 195–222. Amsterdam, Philadelphia: John Benjamins.

Bardel, Camilla & Ylva Falk. 2007. The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research* 23. 459–484.

Bardel, Camilla & Ylva Falk. 2012. The L2 status factor and the declarative/procedural distinction. In Jennifer Cabrelli Amaro, Suzanne Flynn & Jason Rothman (eds.), *Third language acquisition in adulthood* (Studies in Bilingualism 46), pp. 61–78. Amsterdam: John Benjamins.

Bardel, Camilla & Laura Sánchez. 2017. The L2 status factor hypothesis revisited: The role of metalinguistic knowledge, working memory, attention and noticing in third language learning. In Tanja Angelovska & Angela Hahn (eds.), *L3 Syntactic Transfer: Models, new developments and implications* (Bilingual Processing and Acquisition 5), pp. 85–101. Amsterdam, Philadelphia: John Benjamins.

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

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

Berkes, Éva & Suzanne Flynn. 2012. Further evidence in support of the cumulative-enhancement model: CP structure development. In Jennifer Cabrelli Amaro, Suzanne Flynn & Jason Rothman (eds.), *Third language acquisition in adulthood* (Studies in Bilingualism 46), pp. 143–164. Amsterdam: John Benjamins.

Birdsong, David. 2006. Dominance, proficiency, and second language grammatical processing. *Applied PsychoLinguistics* 27(1). 46–49.

Birdsong, David, Libby M. Gertken & Mark Amengual. 2012. *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. COERLL, The University of Texas at Austin. Web. 20 Jan. 2018. Available at: https://sites.la.utexas.edu/bilingual/.

Bowles, Melissa A. 2011. Measuring implicit and explicit linguistic knowledge. What can heritage language learners contribute?. *Studies in Second Language Acquisition* 33. 247–271.

Cabrelli Amaro, Jennifer, João Felipe Amaro & Jason Rothman. 2015. The relationship between L3 transfer and structural similarity across development: Raising across an experiencer in Brazilian Portuguese. In Hagen Peukert (ed.), *Transfer effects in multilingual language development*, 21–52. Amsterdam, Philadelphia: John Benjamins.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT.

Cunnings, Ian. 2017. Interference in native and non-native sentence processing. *Bilingualism: Language and Cognition* 20(4). 712–721.

De Angelis, Gessica. 2007. *Third or additional language acquisition*. Clevedon: Multilingual Matters. https://doi.org/10.21832/9781847690050.

De Vincenzi, Marica, Remo Job, Rosalia Di Matteo, Alessandro Angrilli, Barbara Penolazzi, Laura Ciccarelli & Francesco Vespignani. 2003. Differences in the perception and time course of syntactic and semantic violations. *Brain and Language* 85. 280–296.

Ditman, Tali, Phillip J. Holcomb & Gina R. Kuperberg. 2007. An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology* 44. 927–935.

Ellis, Rod. 2005. Measuring implicit and explicit knowledge of a second language. A psychometric study. *Studies in Second Language Acquisition* 27. 141–172.

Falk, Ylva & Camilla Bardel. 2010. The study of the role of the background languages in third language acquisition. The state of the art. *International Review of Applied Linguistics* 48(2–3). 185–219.

Falk, Ylva & Camilla Bardel. 2011. Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research* 27. 59–82.

Falk, Ylva, Christina Lindqvist & Camilla Bardel. 2015. The role of L1 explicit metalinguistic knowledge in L3 oral production at the initial state. *Bilingualism: Language and Cognition* 18(2). 227–235.

Fallah, Nader & Ali Akbar Jabbari. 2016. L3 acquisition of English attributive adjectives: Dominant language of communication matters for syntactic cross-linguistic influence. *Linguistic Approaches to Bilingualism* 8(2). 1–24.

Fallah, Nader & Ali Akbar Jabbari. 2018. L3 acquisition of English attributive adjectives dominant language of communication matters for syntactic cross-linguistic influence. *Linguistic Approaches to Bilingualism* 8. 193–216.

Flynn, Suzanne, Claire Foley & Inna Vinnitskaya. 2004. The Cumulative-Enhancement Model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism* 1(1). 2–16.

Fox, John & Sanford Weisberg. 2019. *An R companion to applied regression*, 3d edn. Thousand Oaks: Sage Publishing.

Gertken, Libby M., Mark Amengual & David Birdsong. 2014. Assessing language dominance with the Bilingual language profile. In Pascale Leclercq, Amanda Edmonds & Heather Hilton (eds.), *Measuring L2 proficiency: Perspectives from SLA*, pp. 208–225. Bristol: Multilingual Matters.

Giancaspro, David, Becky Halloran & Mike Iverson. 2015. Examining L3 transfer: The acquisition of differential object marking in L3 Brazilian Portuguese. *Bilingualism: Language and Cognition* 18. 191–207.

Godfroid, Aline, Shawn Loewen, Sehoon Jung, Ji-Hyun Park, Susan Gass & Rod Ellis. 2015. Timed and untimed grammaticality judgements measure distinct types of knowledge. Evidence from eye-movement patterns. *Studies in Second Language Acquisition* 37. 269–297.

González Alonso, Jorge González & Jason Rothman. 2017a. Coming of age in L3 initial stages transfer models: Deriving developmental predictions and calling for theory conservatism. *International Journal of Bilingualism* 21(6). 683–697.

González Alonso, Jorge González & Jason Rothman. 2017b. From theory to practice in multilingualism: What theoretical research implies for third language learning. In T. Angelovska & A. Hahn (Eds.), *L3 Syntactic Transfer: Models, new developments and implications* (Bilingual Processing and Acquisition 5), pp. 277–298. Amsterdam/ Philadelphia: John Benjamins.

Haider, Hubert. 2010. *The syntax of German*. Cambridge: CUP.

Harris, Catherine L., Jean Berko Gleason & Ayse Aycicegi. 2006. When is a first language more emotional?. Psychophysiological evidence from bilingual speakers. In Aneta Pavlenko (ed.), *Bilingual minds: emotional experience, expression, and representation*, 257–283. Clevedon: Multilingual Matters.

Heredia, Roberto. 1997. Bilingual memory and hierarchical models: A case for language dominance. *Current Directions in Psychological Science* 6. 34–39.

Hermas, Abdelkader. 2010. Language acquisition as computational resetting: Verb movement in L3 initial state. *International Journal of Multilingualism* 7. 343–362.

Hermas, Abdelkader. 2014a. Multilingual transfer: L1 morphosyntax in L3 English. *International Journal of Language Studies* 8(2). 1–24.

Hermas, Abdelkader. 2014b. Restrictive relatives in L3 English: L1 transfer and ultimate attainment convergence. *Australian Journal of Linguistics* 34. 361–387.

Holmberg, Anders. 2012. Verb second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – An international handbook of contemporary syntactic research*, pp. 343–384. Berlin: Walter de Gruyter.

Hopp, Holger. 2018. Cross-linguistic influence in the child third language acquisition of grammar: Sentence comprehension and production among Turkish-German and German learners of English. *International Journal of Bilingualism* 23(2). 567–583..

Imaz Agirre, Ainara & María del Pilar García Mayo. 2017. Transfer effects in the acquisition of double object constructions in English as an L3. In Tanja Angelovska & Anjela Hahn (eds.), *L3 Syntactic Transfer: Models, new developments and implications* (Bilingual Processing and Acquisition 5), pp. 143–171. Amsterdam/Philadelphia: John Benjamins.

Ionin, Tania. 2012. Formal theory-based methodologies. In Alson Mackey & Susan M. Gass (eds.), *Research methods in second language acquisition*, 30–52. Chichester, England: Wiley-Blackwell.

Jaensch, Carol. 2009. L3 enhanced feature sensitivity as a result of higher proficiency in the L2. In Yan-kit Ingrid Leung (ed.), *Third language acquisition and Universal Grammar*, pp. 115–143. Bristol: Multilingual Matters.

Jegerski, Jill. 2014. Self-paced reading. In Jill Jegerski & Bill VanPatten (eds.), Research methods in second language psycholinguistics, pp. 20–49. New York: Routledge.
Jiang, Nan, Eugenia Novokshanova, Kyoko Masuda & Xin Wang. 2011. Morphological congruency and the acquisition of L2 morphemes. *Language Learning* 61(3). 940–967.
Jin, Fufen 2009. Third Language acquisition of Norwegian objects: Interlanguage transferor L1 influence? In Yan-kit Ingrid Leung (ed.), *Third language acquisition and universal grammar*, 144–161. Clevedon, UK: Multilingual Matters.
Juffs, Alan & Guillermo A. Rodríguez. 2015. *Second language sentence processing*. New York: Routledge.
Just, Marcel A., Patricia A. Carpenter & Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111. 228–238.
Keating, Gregory & Jill Jegerski. 2015. Experimental designs in sentence processing research. A methodological review and user's guide. *Studies in Second Language Acquisition* 37. 1–32.
Kroll, Judith F., Chip Gerfen & Paola E. Dussias. 2008. Laboratory designs and paradigms: Words, sounds, and sentences. In Li Wei & Melissa G. Moyer (eds.), *The Blackwell guide to research methods in bilingualism and multilingualism*, pp. 108–131. Oxford: Blackwell.
Lenth, Russell. 2020. Emmean: Estimated marginal means, aka least-squares means. R Package Version 1.4.8. https://cran.r-project.org/web/packages/emmeans/emmeans.pdf (accessed 24 July 2020).
Lloyd-Smith, Anika, Henrik Gyllstad, Tanja Kupisch & Stefano Quaglia. 2018. Heritage language proficiency does not predict syntactic CLI into L3 English. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2018.1472208.
Lüdecke, Daniel. 2020. *ggeffect. Marginal effects and estimated marginal means from regression Models*. https://www.rdocumentation.org/packages/ggeffects/versions/0.15.1/topics/ggeffect (accessed 24. July 2020).
Marinis, Theodore. 2003. Psycholinguistic techniques in second language acquisition research. *Second Language Research* 19(2). 144–161.
Marsden, Emma J., Sophie Thompson & Luke Plonsky. 2017. A methodological synthesis of self-paced reading in second language research. *Applied PsychoLinguistics* 39(5). 861–904.
McDonough, Kim & Pavel Trofimovich. 2012. How to use psycholinguistic methodologies for comprehension and production. In Alison Mackey & Susan M. Gass (eds.) *Research methods in second language acquisition*, 117–138. Chichester, England: Wiley-Blackwell.
Mykhaylyk, Roksolana, Natalia Mitrofanova, Yulia Rodina & Marit Westergaard. 2015. The Linguistic Proximity Model: The case of verb-second revisited. In Elisabeth Grillo & Kyle Jepson (eds.), *Proceedings of BUCLD 39*, pp. 337–349. Somerville, MA: Cascadilla Press.
Navés, Teresa, Immaculada Miralpeix & María Luz Celaya. 2005. Who transfers more … and what? Crosslinguistic influence in relation to school grade and language dominance in EFL. *International Journal of Multilingualism* 2(2). 113–134.
Na Ranong, Sirirat & Yan-kit Ingrid Leung. 2009. Null Objects in Thai-L2 English-L3 Chinese: An empiricist take on a theoretical problem. In Yan-kit Ingrid Leung (ed.), *Third language acquisition and universal grammar*, pp. 161–191. Clevedon, UK: Multilingual Matters.
Odlin, Terence. 1989. *Language transfer. Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
Oxford Placement Test. https://www.oxfordonlineenglish.com/english-level-test (accessed 24. February 2020).
Paradis, Michel. 2004. *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Peirce, Jonathan W. 2007. PsychoPy – psychophysics software in Python. *Journal of Neuroscience Methods* 162. 8–13.

Puig-Mayenco, Eloi, David Miller & Jason Rothman. 2018. Language dominance and transfer selection in L3 acquisition. In Jacee Cho, Michael Iverson, Tiffany Judy, Tania Leal & Elena Shimanskaya (eds.), *Meaning and Structure in Second Language Acquisition*. Studies in Bilingualism (55), 229–260. Amsterdam: John Benjamins.

Puig-Mayenco, Eloi, Jason Rothman. & Susagna Tubau. 2020. Language dominance in the previously acquired languages modulates rate of third language (L3) development over time: A longitudinal study, *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2020.1792408.

Rah, Anne. 2010. Transfer in L3 sentence processing: Evidence from relative clause attachment ambiguities. *International Journal of Multilingualism* 7(2). 147–161.

Ringbom, Håkan. 1986. Crosslinguistic influence and the foreign language learning process. In Eric Kellerman & Michael. Sharwood-Smith (eds.), *Crosslinguistic influence in second language acquisition*, pp. 150–162. Oxford: Pergamon Press.

Roberts, Leah & Sarah Liszka. 2013. Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research* 29(4). 413–439.

Rothman, Jason. 2010. On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Teaching* 48. 245–273.

Rothman, Jason. 2011. L3 syntactic transfer selectivity and typological determinacy: The Typological Primacy Model. *Second Language Research* 27. 107–127.

Rothman, Jason. 2013. Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In Sergio Baauw, Frank Dirjkoningen & Manuela Pinto (eds.), *Romance languages and linguistic theory*, pp. 217–247. Amsterdam: Benjamins.

Rothman, Jason. 2015. Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition* 18. 1–12.

Rothman, Jason & Jennifer Cabrelli Amaro. 2010. What variables condition syntactic transfer?. A look at the L3 initial state. *Second Language Research* 26. 189–218.

Rothman, Jason, José Alemán Bañón & Jorge González Alonso. 2015. Neurolinguistic measures of typological effects in multilingual transfer: Introducing an ERP methodology. *Frontiers in Psychology* 6. 1087.

Rothman, Jason, Jorge González Alonso & Eloi Puig-Mayenco. 2018. *Third Language Acquisition and Linguistic Transfer*. Cambridge, UK: Cambridge University Press.

Sağın Şimşek, Çiğdem. 2006. *Third language acquisition: Turkish-German bilingual students' acquisition of English word order in a German educational setting*. Münster: Waxmann.

Schad, Daniel, Shravan Vasishth, Reinhold Kliegl & Shihao Du. 2020. *How to capitalize on a priori contrasts in linear (mixed) models: A tutorial*. osf.io/7ukf6 (accessed 14 August 2020).

Schütze, Carson T. 2016. *The empirical base of linguistics: Grammaticality judgments and linguistics methodology*. Berlin: Language Science Press.

Schwartz, Bonnie D. & Rex Sprouse. 1996. L2 cognitive states and the full transfer/full access model. *Second Language Research* 12. 40–72.

Slabakova, Roumyana. 2017. The scalpel model of third language acquisition. *International Journal of Bilingualism* 21. 651–665.

Tokowicz, Natasha & Tessa Warren. 2010. Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study. *European Journal of Cognitive Psychology* 22(7). 1092–1106.

Tokowicz, Natasha, Erica B. Michael & Judith F. Kroll. 2004. The roles of study-abroad experience and working memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition* 7. 255–272.

Tremblay, Annie. 2005. Theoretical and methodological perspectives on the use of grammaticality judgment tasks in linguistic theory. *Second Language Studies* 24(1). 129–167. http://hdl.handle.net/10125/40679.

Westergaard, Marit. 2009. *The acquisition of word order. Micro-cues, information structure, and economy* (Linguistik Aktuell/Linguistics Today 145). Amsterdam: John Benjamins.

Westergaard, Marit, Natalia Mitrofanova, Roksolana Mykhaylyk & Yulia Rodina. 2017. Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21. 666–682.

Williams, Sarah & Bjorn Hammarberg. 1998. Language switches in L3 production: Implications for a polyglot speaking model. *Applied Linguistics* 19(3). 295–333.

Wickham, Hadley. 2020. gplot2: Elegant graphics for data analysis. https://ggplot2-book.org/index.html (accessed 24 July 2020).