WHERE SYNTACTIC INTERFERENCE PERSISTS: THE CASE OF HEBREW WRITTEN BY NATIVE ARABIC SPEAKERS

EIHAB ABU-RABIAH[1], ROEY J. GAFTER[2] & RONI HENKIN[2]

1. University of Utah
2. Ben-Gurion University of the Negev

Abstract
In this paper we analyze the acquisition of Hebrew syntax in a corpus of essays, written by 22 native speakers of Arabic after studying Hebrew for a decade. Each of the participants contributed two essays to the corpus: one when they were in the 11th grade of high school, and a second essay a year later. We categorized the syntactic errors, and explored the relationship between persistence, interference and developmental errors. Statistical analysis showed interference to be involved in the vast majority of the errors that persisted most between the two time-points, whereas almost all the improvement over the year was in developmental errors with no interference. These results contradict a common claim that interference, initially predominant, decreases over time with relation to developmental errors. We propose that a key difference is that much of the theorization in Second Language Acquisition (SLA) is based on findings in unrelated language dyads, with English as L1 or L2, whereas the languages in our study are closely related, yet differ considerably in their syntax. We conclude that more research on syntactic interference in the acquisition of related languages is necessary in order to reveal findings diverging from many typically attested patterns.

Keywords: second language acquisition; interlanguage; interference; transfer; developmental effects; stabilization; persistence

Abu-Rabiah, E., Gafter, R.J., & Henkin, R. (2023). Where syntactic interference persists: the case of Hebrew written by native Arabic speakers. L1-Educational Studies in Language and Literature, 22, 1-20. https://doi.org/10.21248/l1esll.2023.22.1.399
Corresponding author: Eihab Abu-Rabiah, Department of Linguistics, The University of Utah, Salt Lake City, UT 84112, USA, email: aburabiah.e@utah.edu

1. INTRODUCTION

Much of the theorization of Second Language Acquisition (SLA) is based on studies of English as L2 or, to a lesser extent, as L1 (Granger, 2015, p. 9). Furthermore, the L1 and L2 under research are usually unrelated. As an outcome of this bias in the field, certain attested patterns which are thought to be general or "typical" may in fact turn out to be more limited than assumed, when a broader set of language dyads is considered.

In this paper we argue that this research gap has important consequences for understanding the role of transfer, that is, the degree to which a language learner's L1 affects and shapes their learning process. Transfer which hinders acquisition of L2 features differing from their parallels in the learner's L1, also known as 'interference', is not uniform across the learning trajectory. The common view is that interference is expected to play less of a role in advanced stages of learning (Ellis, 1985, pp. 24–25; Heydari & Bagheri, 2012, p. 1588). With respect to the different language subsystems, it has been claimed that interference in sentence production and morphosyntax is limited compared to its role in the acquisition of phonology or the lexicon (MacWhinney, 2005, p. 60). Taking these two claims together, syntactic interference is expected to be minimal among advanced learners. In this paper, we demonstrate that this is not always the case, by examining the persistence of syntactic errors among advanced learners of Hebrew who are native speakers of Arabic. Crucially, unlike in much of the previous research, the languages in question are two genetically related languages that nevertheless differ considerably in their syntax.

Our data come from a corpus of Hebrew writing by high school students in an Arabic-medium high school in the Negev, southern Israel, which was collected at two time-points separated by a year: when the students were in the 11th and 12th grade. We show that in our corpus an overwhelming number of persisting errors can be attributed to interference from Arabic. While these results may be surprising in light of previous research, we propose that they can be understood given how the particular pair of languages related to each other. Therefore, we argue that our data highlight the need for more research on a more diverse set of language dyads when investigating the acquisition of L2 syntax.

2. THEORETICAL BACKGROUND

2.1 Interference in SLA research

In assessing how L2 learners' interlanguages differ from those of L1 speakers, a broad distinction can be made between two types of errors. The first is interference errors, which are the outcome of applying L1 features to L2. This type contrasts with errors that are not attributed to the learner's L1, but rather to developmental cognitive

strategies, as found in L1 acquisition (Dulay & Burt, 1974, p. 129). These are also referred to as 'general L2 processing' and 'general learner effects' (Odlin, 2003, p. 472 ff.; Roberts et al., 2008).

An ongoing debate concerns the relative impact of these two factors in SLA, with opinions often dependent on the contemporary SLA theory, as it developed over the years. At Phase 1 in Rod Ellis's (2021) summary of SLA history, earlier behaviorist accounts, which credited transfer as responsible for all L2 learning, were replaced since the 1960s by generative or mentalist accounts, whereby language learning was recognized as a 'creative construction process' not merely 'habit formation'. SLA was consequently theorized using the same mechanisms as L1 acquisition research (MacWhinney, 1992, p. 5). Under this approach, Error Analysis (EA) in teaching should not involve L1 at all (Heydari & Bagheri, 2012, p. 1583), in contrast with the former approach that had based all SLA on Contrastive Analysis (CA) of the languages concerned. Dulay and Burt (1974) studied English L2 speech of children speaking Spanish, Japanese, Chinese, and Norwegian, and found interference from L1 to account for just 3–5% of the errors, whereas developmental phenomena, characteristic of L1-acquisition, rated at 85–87%. Based on the small number of interference-related errors, their pedagogical suggestions were that 'less explicit teaching of ESL [English as a second language] syntax to children may produce better learning' (Dulay & Burt, 1974, p. 129).

The scientific debate on the relative importance of interference and developmental factors in causing SLA errors continues to this day. Heydari and Bagheri (2012) present findings supporting both sides, based mainly on case studies of English as a foreign language (EFL). Although general agreement was not found, a process seems to be agreed upon in CA and EA research, whereby interference prevails more in the first stages of acquisition, especially in a classroom context, whereas developmental errors gain ground later (Ellis, 1985, pp. 24–25, 37), when "more and more intralingual errors are manifested" (Heydari & Bagheri, 2012, p. 1588). However, although intuitively expected to decrease over time, this is not always the case with transfer: "transfer effects... may fluctuate and even increase with advancing L2 proficiency" (Jarvis, 2015, p. 23).

A major factor affecting the degree of transfer is the similarity of the languages concerned. Similarity promotes positive transfer, so is beneficial on the one hand (Jarvis, 2015, p. 30); but on the other hand, it also promotes interference. The Crucial Similarity Measure (Wode, 1978, pp. 102, 116) is based on the finding that interference is most likely in cases of 'relative similarity' between the languages: not total identicalness of L1 and L2, but not extreme dissimilarity either, as the inclination to transfer decreases in such cases, as when an English-speaker is acquiring Chinese (Ellis, 1985, pp. 34–35). Notably, transfer is caused not so much by objective similarities and differences between the languages per se, but rather by the learners' perception of these (Jarvis, 2015, p. 23) or what Weinreich (1968, p. 7) calls 'interlingual identifications'. For example, Diab (1996, p. 82) noted that Lebanese students erred

most in English when they perceived the corresponding structures in the two languages as similar.

2.2 Interference in syntactic domains

Transfer, both positive and negative, occurs in all subsystems of language, but the relative strength of interference in the diverse linguistic domains is not unanimously agreed upon. Transfer has been recognized as considerable in subsystems such as lexicon (Roberts et al., 2008, p. 334) and phonology (Ellis, 1985, pp. 19, 40), whereas "L1 transfer in the areas of sentence production and morphosyntax is limited by the fact that morphosyntax is the most language-specific part of a target language. Because the mappings are hard to make, transfer in this area is minimized" (MacWhinney, 2005, p. 60). Odlin (2003, p. 437) notes that "Some... have been skeptical about transfer in syntax and morphology, but such skepticism is unwarranted". Positive transfer has been found, particularly between languages that are similar, even in issues where influence has been previously denied, such as tense and aspect. This influence increases when grammatical morphemes have similar equivalents in the two languages, whereas differences in these issues cause more difficulties (Jarvis, 2015, pp. 24–25).

Within syntax, domains that have been found to be particularly difficult for EFL learners from diverse L1 backgrounds include use of the article, the verb, and prepositions (Heydari & Bagheri, 2012, p. 1585). Of these, the domain that will concern us most is prepositions, which are widely recognized as problematic in SLA, and prone to interference, also resistant to corrective feedback or "less treatable" than other domains (Ferris, 1999; Bitchener et al., 2005). The difficulty with prepositions is largely due to cross-linguistic arbitrariness in their use, arbitrariness being one of Todeva's three high-risk categories of features prone to stabilization (Long, 2003, p. 518).

2.3 Transfer between Arabic and Hebrew

The research surveyed above has evaluated the relative prevalence of interference and developmental errors in general and for specific linguistic domains of L2. It has, however, concentrated mainly on English as L2 and most of the L1s studied were rather distant, genetically and typologically, including Taiwanese, Korean, Japanese, French, Italian, Norwegian, Iranian, and Arabic. Such conditions of relative distance, under the Crucial Similarity Measure, do not encourage interference. The L1 and L2 in our study, however, are two genetically and typologically related Semitic languages, with many similarities but also many differences, and coexisting in an asymmetric social relation (Henkin-Roitfarb, 2011).

As two Semitic 'sister' languages, Hebrew and Arabic are genetically and typologically close, though not mutually intelligible. They share basic root/pattern morphology in both inflection and derivation; a common ancient lexical inventory is evident

in numerous lexical cognates. However, they have a totally different script; more importantly, due to its unique history as a language revitalized for speech purposes, Modern Hebrew has many non-Semitic influences as well. The temporal system, for example, is quite different from that of Arabic: while Arabic has preserved the ancient binary aspectual system of perfect and imperfect, Modern Hebrew has a temporal system of three tenses: past, present, future. Furthermore, Hebrew has a system of differential object (accusative) marking, which is problematic for speakers of many other languages, including Arabic.

Due to the basic similarity and close affinity between these two languages, mutual acquisition, in terms of comprehension and production, is expected to proceed more smoothly and quickly than in the case of distant languages (Jarvis, 2015, p. 30), according to SLA research literature whereby "learners of a related language regularly outperform learners of an unrelated language" since the former "have a far smaller learning burden than learners of a distant language" (Ringbom & Jarvis, 2009, pp. 112, 115).

Nevertheless, the previous research on the acquisition of Hebrew by Arabic speakers does not present a clear picture on how dominant the role of interference is. Most of the earlier work tended to list errors and identify L1 influence in a contrastive error analysis approach that predominated SLA research from the 1950s to the 1970s (Hinkel, 2005, p. 615). In this vein, Bassal (2007) dedicates half of his article to differentiating interference from errors due to difficulties in the target language, and the other half to contrastive analysis and ways of exploiting the similarities in instruction. In contrast, Shehade (1998) warns that the extreme closeness may well be detrimental in teaching Hebrew to Arabic speakers and vice versa. Research that looks into possible interference from Arabic in L2 syntax, and which is therefore of special relevance to our study, is no exception: some researchers find a predominance of interference effects (Noor, 199, p. 1461; Al-Khresheh, 2011, p. 426), while others see interference as minor and secondary (AbiSamra, 2003, p. 22; Bataineh, 2005, p. 56).

Recent experimental work on domain differential interference monitoring in different-script languages (Prior et al., 2017) focuses on the Hebrew of Arabic-speaking university students, whose proficiency in Hebrew was assessed by standard entrance exams. Activation and inhibition of L1 interference and its correlation with proficiency were tested by lexical and syntactic judgment tasks, involving a wide range of syntactic structures that differ in Modern Hebrew and Arabic. Most relevant to the present study is the conclusion that syntactic interference, but not lexical interference, was reduced with rising L2 proficiency. Conversely, Haskel-Shaham et al. (2018), studying the academic writing of Palestinian Arabic-speaking teachers of Hebrew, found persistence of errors to decrease primarily in discourse, then in the lexicon, and least of all in grammar (phonology and syntax).

Most recently, Henkin (2020) found support for a multiple effect approach to persistence. Localized errors of phonology, orthography, and morphology generally declined faster than syntactic errors, which persisted especially as preferences in

structures that occur in L1 Hebrew, but are marked for discourse-pragmatic effects. The main factor proposed to account for the persistence differences was what she called 'error salience', the likelihood of an error to be detected by a teacher, working under time pressure. Error salience was found to be maximal in localized features, which are subject to rigid grammatical rules, whereas preferences and less localized features of syntax were less error-salient and most persistent.

3. PRESENT STUDY

The present study is based on a corpus of essays in Hebrew written by a group of Hebrew learners: high school students who are native speakers of Negev Arabic. Each participant contributed two essays at two different points in time: first, when the participants were in the 11th grade (ages 17–18), and then a year later, when they were in the 12th grade. Following the tradition of Contrastive Interlanguage Analysis (CIA) on a learner corpus (Granger, 2015), we identified all syntactic errors in the corpus, in order to assess the persistence of interference phenomena as against developmental phenomena in the syntax of their written Hebrew.

3.1 Aims and research questions

Using our corpus of Hebrew written by Negev Arabic-speaking high school students, our aim is to establish the relations between interference and developmental phenomena in the persistence of syntactic errors in a pair of closely related languages. Thus we have three interrelated research questions. First, we wish to establish if there is a general decrease in syntactic errors over time. Our initial prediction is that since the students have a year of learning between the two sections of the corpus, we do expect such a decrease. Our second research question is—assuming that rates of syntactic errors indeed do decrease—how do rates of decrease in interference errors compare with those of developmental errors with no interference? In keeping with the research literature summarized in section 2.1, we expect decrease in interference errors to surpass the decrease in developmental errors. Finally, we wish to examine if there are particular domains in each of the two categories that persist more than others. Based on the review in section 2.2, we expect preposition errors to persist more than errors in other syntactic domains.

3.2 Participants

The participants in our sample are members of the Bedouin community native to the Negev, in southern Israel. The community language is Negev Arabic (Henkin, 2010; Shawarbah, 2012), which may be considered a Bedouin variety of Palestinian Arabic. In Arabic-language medium schools in Israel, Hebrew is officially introduced at 2nd grade (ages 8–9). It is taught for 3–5 hours/week throughout the school years, culminating in the final exam in Hebrew as L2 for the Arab sector. In high school the

language of instruction in Hebrew classes is officially Hebrew, but in practice it tends to lapse to spoken Arabic, the L1 of almost all the teachers.

The sample originally included 23 participants, but one had to be excluded since the written product was completely incomprehensible, and we could not discern what syntactic structures the sentences were intended to represent. Therefore, the sample consists of 22 participants (nine male, thirteen female) all native speakers of Negev Arabic; they all study in the same advanced level science-track class, at a public high school in the Northern Negev. They all started instruction at the same age (2nd grade), and have gone through a similar learning trajectory, and therefore have received the same amount of formal exposure to Hebrew (though there may be differences in the amount of non-formal exposure). In the year that passed between the two time-points, the participants, who were all in 11th grade, studied 3 weekly hours of Hebrew language and literature. Their Hebrew teacher during the one-year study period was also a native speaker of Negev Arabic. The Hebrew language curriculum in this grade focuses on verbal morphology; in the domain of syntax, the main topic is how to construct and differentiate different sentence types. Thus, the students had relatively little explicit instruction on the syntactic issues examined here, with the possible exception of 'complex sentences'.

3.3 Methods

The participants watched an animated video called "Bridge", created by Ting Chian Tey. The silent video depicts two animals approaching each other from two sides of a narrow bridge, and discovering that they need to negotiate a compromise in order to cross. The video was given to help the students start the complex process of writing without providing any linguistic help. It lasted less than three minutes, during which no discussion was conducted. Afterward, the students were asked to write an argumentative essay dealing with the question: "Are you willing to negotiate and reach a compromise in order to promote a common goal?". The task was performed under supervision, with the teacher ensuring that no one received any help. All students completed their writing in less than 90 minutes (two consecutive lessons). The task was first performed in 11th grade. The essays were not returned to the writers, and no feedback was given. The procedure was repeated a year later in 12th grade.

3.4 Data coding and analysis procedures

As a first step, we extracted all the syntactic errors from the corpus, using mainstream native Hebrew as the reference language variety (RLV, Granger, 2015, 8). The first author of the paper (a native Arabic speaker with near-native Hebrew and a PhD in Hebrew linguistics) and the third author (a native Hebrew speaker with a PhD in Arabic linguistics) each separately went over the entire corpus and marked every place in which the syntax deviated from the RLV as a putative error. They then compared their error coding, which showed a 90% rate of inter-rater agreement. In the

cases in which there was no initial agreement, the nature of the error was discussed with the second author of the paper (a native Hebrew speaker with a PhD in linguistics) until a unanimous agreement was achieved between all three.

The authors then analyzed the syntactic errors by sorting them into six primary syntactic domains, which we will introduce below. Furthermore, errors were analyzed as pertaining to one of two types: first, interference errors, i.e., those that can be explained by the syntactic structure of the writers' L1 (Arabic) as at least one source of influence, though not necessarily the only source. Second, developmental or general learner errors without interference, i.e., those that cannot be explained by influence from the learners' L1. These classifications were based on SLA research literature, their own research on Hebrew as an L2, and their experience in teaching academic Hebrew and Hebrew as an L2.

Below we show two examples from the corpus for each of the six domains: one which shows interference, followed by a developmental error with no interference. Note that in isolation, some of the examples would be grammatical in L1 Hebrew as well, but the context of the essay made it clear that the intended meaning did not match the forms produced.

a. Agreement:

(1) ve- kax ha- beʕayot tatḥil
and so the problem.F.PL begin.FUT.F.3SG
Intended meaning: 'and so the problems will begin'

(2) yad eḥad
hand.F one.M
Intended meaning: 'one hand'

In example (1) the subject and the verb are both expected to be in the plural form. The lack of agreement here matches Arabic agreement patterns, by which an inanimate plural noun takes feminine singular agreement markers. The agreement mismatch is therefore attributed to interference. In (2), the agreement mismatch (that is, the use of a masculine number to refer to a feminine noun) cannot be due to interference from Arabic, since Arabic would also have gender agreement in the equivalent sentence, and 'hand' is feminine in both languages.

b. Tenses:

(3) ani eḥšov
I think.FUT.1SG
Intended meaning: 'I think'

(4) i.efšar še- šte mexoniot laʕavor
Impossible that two car.PL pass.INF
Intended meaning: 'It is impossible for two cars to pass'

In Modern Hebrew, there are three main tenses: past, present, and future, while Arabic has only two: past and non-past, which can express either present or future. Example (3) means "I will think" but from the context it is clear that the student was referring to the present. Hence, interference from Arabic may underlie the mismatch in tenses. Example (4), however, shows a use of the infinitive form where Hebrew would expect an inflected verb—this cannot be attributed to interference from Arabic, which would also use an inflected verb there, and in fact has no infinitival form.

c. Definiteness:

(5) ze ha- davar
this the thing
Intended meaning: 'this thing'

(6) me- ḥanut ha- krova
from store the nearby
Intended meaning: 'from the nearby store'

In Modern Hebrew, the attributive demonstrative pronoun usually follows the head noun and each is preceded by the determiner *ha-*. Example (5) shows interference from Arabic where a non-determined demonstrative pronoun precedes the determined noun. In (6), the determiner *ha-* was expected to precede both the noun 'store' and the adjective 'nearby', but it appears only before the adjective. The omission of the compulsory determiner cannot be due to interference, because both adjective and noun would be preceded by a determiner in Arabic as well.

d. Governed prepositions:

(7) halaxti ʕal ha- ḥaver šeli
walk.PS.1SG on the friend my
Intended meaning: 'I walked over to my friend'

(8) levater davar
forgo.INF thing
Intended meaning: 'to forgo something'

The Hebrew preposition *ʕal* 'on' is often used by Arabic speakers in the sense of 'to (a goal)', since in Arabic *ʕal* is governed by 'go'. This is shown in Example (7), in which the preposition *le-* 'to' would be expected. In (8) there is no preposition, although the Hebrew verb *levater* ('forgo') requires *ʕal*. Omission of the preposition here is not due to interference, since the governed preposition is obligatory in both languages in this case.

e. Complex sentences:

(9) ḥalom ata roce lehagšim
dream you want.PR.M.SG fulfil.INF

Intended meaning: 'a dream that you want to fulfil'

(10) im še- ze ha- štayim
If that this the two
Intended meaning: 'if these two'

In (9), the relative clause has no subordinating conjunction, whereas in Modern Hebrew the complementizer *še-* ('that') is obligatory in this case. The lack of *še-* matches Arabic syntax, in which asyndetic relative clauses are the rule if the head noun is indefinite. Example (10) has a superfluous occurrence of *še-*, which cannot be attributed to Arabic, in which this structure would be wrong as well.

f. Pronouns and copulas:

(11) ani avater im roe
I forgo.FUT.1SG if see.PR.M.SG
Intended meaning: 'I will forgo if I see'

(12) yeš anašim še- hem mitʕarvim
there.are people that they interfere.PR.M.PL
Intended meaning: 'there are people who interfere'

Arabic and Hebrew differ in the extent to which they permit null subjects. In (11), the subordinate clause ("if I see") is lacking a first-person pronoun, which would be obligatory in Hebrew, but not in Arabic. In (12), the pronoun *hem* 'they' is as superfluous in Hebrew as it is in Arabic, so this error is not a case of interference.
We excluded errors in the gray area between syntax and vocabulary, such as using a certain conjunction with the meaning and complementation of another conjunction. An example is using an infinitive, instead of a clause, after Hebrew *ki* 'because', due to Arabic *kay* 'in order to'.

It is important to note that the way we operationalized "errors" makes a distinction between errors and preferences. The latter are often formulated in SLA literature as over- and underuses. Both are "features of non-nativeness of the interlanguage, i.e., not only errors [...], but also over- and underuses, which represent a (positive or negative) deviation from a certain statistical norm that characterizes native performance" (Gilquin, 2001, p. 99). While errors deviate from the language norms of native speakers, who show no variation in these features, non-native preferences are just quantitatively non-representative of the majority of speakers, contexts, and styles. Accordingly, in comparing the learner's corpus to the RLV, our aim was to include only errors in the analysis, and exclude any deviations from the RLV which could be explained as preference. Nevertheless, the border between these two categories is not always clear-cut. For example, when L1 information structure is transferred from L1 Swedish to L2 German, Bohnacker and Rosén (2007) consider this a case of preferences, not errors, since the resulting forms are correct in the target L2, although their frequency is not native-like. In contrast, underuse of English passives

by Hebrew speakers has been considered pragmatically inaccurate, i.e., erroneous (Seliger, 1989, as presented in MacWhinney, 1992, p. 12). In the case of Arabic-speaking EFL learners, overuse of coordination and parallelism has been noted as most characteristic (Kaplan, 1966; Hinkel, 2005, p. 618).

In our study, we ignored preferences such as overuse of coordination and other L1-characteristics, as long as we judged the structure to be acceptable to native speakers in the given context. However, the same form would be considered an error if rejected as not acceptable in the given context, even if found in other contexts. For example, the Hebrew accusative marker *et,* normally required before a definite direct object, is sometimes omitted by native speakers in certain contexts, registers, and styles, but its omission sounds strange in (13):

(13) ani	šoel	[et]	axi
I	ask.PR.M.SG	[ACC]	my.brother

Intended meaning: "I ask my brother"

The omission of *et* in (13) is thus classified as an error, one that is both interference-related and developmental, since it is common in acquisition of L2-Hebrew, regardless of L1.

4. RESULTS

We identified 711 syntactic errors overall: 386 at the first time-point (TP1) and 325 at the second. That is, there were fewer errors at the second time-point (TP2), as one may expect if students are increasing in proficiency. Importantly, however, the essays were on average somewhat shorter at TP2: the corpus of essays consisted of 3,706 words at TP1, and of 3,087 words at TP2 (6,793 words overall). Thus, the rate of syntactic errors per words remains virtually the same, as shown in Table 1.

Table 1. Number of syntactic errors at each time-point

	TP1	**TP2**	**Total**
Overall number of errors	386	325	711
Overall number of words	3706	3087	6793
Errors per 100 words	10.42	10.53	10.47

Out of the syntactic 711 errors overall, 551 (77%) were coded as interference errors, whereas only 160 (23%) do not involve possible interference. That is, the vast majority of syntactic errors in the corpus are at least compatible with being the results of interference. Table 2 shows the number of errors with and without interference at each of the two time-points.

Notably, at TP1 interference errors accounted for 73.1% of the errors, whereas at TP2 their number stayed relatively constant while the number of non-interference

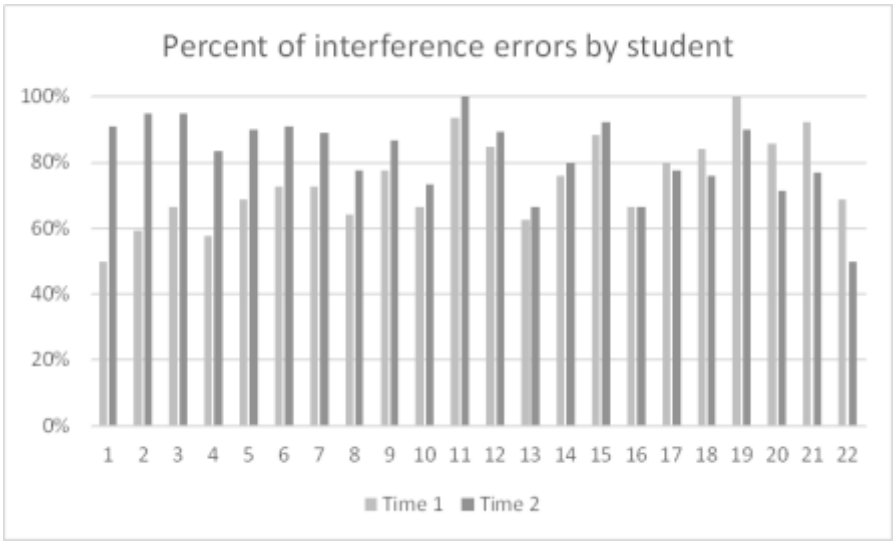
errors was considerably smaller, and thus the interference errors accounted for 82.8% of the errors.

Table 2. Number of syntactic errors with and without interference

	Time-point 1		**Time-point 2**	
	# errors	Errors per 100 words	# errors	Errors per 100 words
+Interference	282	7.61	269	8.71
-Interference	104	2.81	56	1.81
Total	386	10.42	325	10.53

In order to assess the inter-subject variation, we calculated for each student the percentage of interference errors at each time-point. These rates are shown in Figure 1. As the figure shows, for the majority of participants (15 out of the 22 students), the percentage of interference errors at TP2 is, unexpectedly, greater than at TP1.

Figure 1. Percent of syntactic interference errors by student



A paired-samples t-test was conducted in order to compare the two time-points, and it shows that the percentage of interference errors at TP1 (M = 0.74, SD = 0.13) is significantly lower than at TP2 (M= 0.82, SD = 0.12), $t(21) = 2.23$, $p < .05$. (Cohen's d = 0.48, moderate effect size). In other words, to the extent that the students are improving by reducing their errors, this reduction in errors is due almost entirely to those errors that do not involve interference.

In order to further determine which are most likely to persist, we looked at the six most common syntactic error domains, as outlined in the previous sections, resulting in 610 errors. The remaining 101 errors (62 at TP1 and 39 at TP2) were in

various idiosyncratic domains, of which there were too few to generalize, and hence they were not included in this analysis. Table 3 shows the number of errors in each domain at each of the time periods, and its percent of the total errors at that time-point. As the table shows, for each of the six domains analyzed, the relative part of the errors is mostly stable between the two time-points.

Table 3. Number of errors by domain

	TP1		**TP2**	
Governed prepositions	86	22.3%	81	24.9%
Agreement	70	18.1%	51	15.7%
Complex sentences	58	15.0%	53	16.3%
Tenses	50	13.0%	39	12.0%
Pronouns and copulas	34	8.8%	38	11.7%
Definiteness	26	6.7%	24	7.4%
Other	62	16.1%	39	12%
Total	386	100%	325	100%

In Table 4 we split the errors in each of the domains into with or without interference.

Table 4. Number of errors by domain and interference

	With interference				**No interference**			
	TP1		**TP2**		**TP1**		**TP2**	
Governed prepositions	66	17.1%	64	19.7%	20	5.2%	17	5.2%
Agreement	36	9.3%	44	13.5%	34	8.8%	7	2.2%
Complex sentences	57	14.8%	51	15.7%	1	0.3%	2	0.6%
Tenses	43	11.1%	37	11.4%	7	1.8%	2	0.6%
Pronouns and copulas	22	5.7%	25	7.7%	12	3.1%	13	4%
Definiteness	10	2.6%	16	4.9%	16	4.1%	8	2.5%
Other	48	12.4%	32	9.8%	14	3.6%	7	2.2%
Total	282	73.1%	269	82.8%	104	26.9%	56	17.2%

As the tables show, some domains have much higher error ratios than others, but the share of each domain from the overall syntactic errors remains mostly constant over the two time-points. An interesting exception is agreement errors that do not involve interference, which exhibit a massive drop—from 8.8% of all syntactic errors at TP1 to a mere 2.2% at TP2. At the same time, agreement errors that do involve interference rise from 9.3% to 13.5%.

5. DISCUSSION

The first finding is that although the absolute number of syntactic errors slightly decreased between the two time-points, the overall level measured in errors per 100 words remained nearly identical. That in itself, while contradicting our first predic-

tion, is not necessarily surprising, given that in 11th grade a year may not be a sufficiently long period of time in which to improve linguistic competence in one's L2. This length of time was found to be insufficient for significant changes in other language areas, such as vocabulary (Zheng, 2016; Kalantari & Gholami, 2017).

The second finding is that the vast majority of syntactic errors at both time-points are at least compatible with being the results of linguistic interference from the students' native Arabic. Furthermore, we found these errors to be particularly persistent: the small improvement that was observed between the two time-points was due entirely to the number of non-interference errors being reduced, whereas the number of interference errors remained constant (and even grew in terms of errors per 100 words). Due to how we defined 'interference'—as occurring wherever the erroneous structure matches a grammatical Arabic structure—we cannot be certain that interference indeed was at play in all these cases, only that the error was compatible with it. However, the significant difference between the percentage of interference errors at the two time-points shows that the errors that we labeled as 'with interference' do differ in persistence from the others, and therefore the most likely explanation is that indeed interference was the reason for their persistence. This finding contradicts the prediction based on the literature we presented above (e.g., Ellis, 1985), that over time interference errors should decrease whereas developmental errors increase (in their relative proportion of all errors), since rising competence in L2 should increase ability to manage L1-interference, whereas unpredictable irregularities in the target language still create difficulties. We find, however, specifically in the domain of syntax, that interference rises whereas developmental errors drop in the overall ratio of errors.

With respect to the different domains of errors, the data show that some are far more frequent than others. Specifically, errors involving governed prepositions are seen to be the most common, comprising almost a fifth of the entire syntactic error corpus. This is not trivial, as governed prepositions occur only after certain verbs, whereas agreement, for example, is obligatory in both verbal and nominal predication, noun-attribute combinations, and pronouns, so there is much more 'room' for errors. As we saw above (section 2.2) Bitchener et al. (2005) found similar results, namely that prepositions constitute 29.23% of all errors in ESL students' writing. Interestingly, this top rank for syntactic errors in general is retained by prepositions in the context of errors that include interference, where their portion of all syntactic errors ranges from 17% at TP1 to 19% at TP2. Conversely, prepositional errors with no interference are at a low 5%. In the Arabic-Hebrew context, the problem of interference in the use of prepositions is well-known. The literature on the Hebrew interlanguage of Arabic speakers, as summarized in Section 2.3 above, contains many lists of preposition errors under L1 influence by Arabic speakers at all levels of education, including professional teachers of Hebrew (Shehadeh, 1998; Bassal, 2007). Misuse of prepositions was targeted as one of the most persistent errors over two years of Hebrew studies at college (Haskel-Shaham et al., 2018). Henkin (2020) found that this domain shows high error salience, i.e., teachers identify it easily and correct it

more than other errors. She attributed this high error salience, among other factors, to rigidity in prepositional use, whereas domains of a more flexible, stylistic nature, such as word order or zero-subject, are less error salient. Despite this high error salience, however, preposition use was found to be highly persistent in that study, and this is exactly what we now see in our results.

However, when we compare the relative share of each domain from the total errors at each time-point, we see that they remain almost entirely constant. Therefore, there is no evidence that these domains differ in how persistent they are. Specifically, we hypothesized that there would be particularly high persistence of preposition errors that are due to interference. But while our results show that governed prepositions may be particularly *difficult* to acquire, there is no evidence that errors in this domain are any more *persistent* than the other domains.

The only exception, in which the share of the errors greatly decreases between the two time-points, is that of agreement errors that do not involve interference. Indeed, much of the overall improvement in the corpus can be attributed solely to this domain, which does appear to be less persistent than the others. One potential explanation would emerge if agreement were precisely what the students were taught in class during the course of the year. However, as mentioned in section 3.2, the 11th grade curriculum does not focus on that. Of our six domains, the only one which receives considerable explicit instruction is "complex sentences", since the curriculum addresses different sentence types, and how they are formed—and errors of this type were no less persistent. Thus, it does not seem to be the case that the improvement in agreement is tied to its receiving particular attention in class.

The explanation that we propose is that agreement, in both Arabic and Hebrew, is an extremely pervasive feature of the grammar, dominant in nouns, verbs, adjectives, and pronouns. It is acquired very early in L1 and is always present in the linguistic awareness of the speakers, who are attuned to this aspect of grammar. When Arabic speakers acquire Hebrew, an L2 that is very similar to their L1 in this respect, including many similar patterns and forms, we expect basic agreement (excluding irregularities) to be acquired relatively early and 'naturally' due to the built-in focus on this aspect of language. The many errors at TP1 may be due to the pervasiveness of the category of obligatory grammatical agreement as well as to the cognitive load of writing an argumentative essay in L2. By TP2, however, a clear improvement is seen, precisely in the domains that are easier for these students to acquire and that fall into place once the cognitive load is lessened through the general rise in proficiency and familiarity with the task. Interestingly, however, agreement errors due to interference do not decrease. In fact, they increase. This may be due to more complex syntax that emerges at TP2, with more attributes that require agreement, leading to more potential for L1-interference. And we see that while agreement is a grammatical category that comes naturally to these speakers, it is their L1-Arabic agreement that presides, which is often at odds with Hebrew rules.

Thus, taken together, we believe that our results show that interference is by far both the most prominent and the most persistent obstacle in acquisition of Hebrew

syntactic patterns by Negev Arabic-speakers. Such pervasiveness and persistence of interference is unexpected in view of much of the extant SLA research, usually based on L2 English, whereby interference is generally agreed to be more prominent at beginners' level, whereas developmental errors gain ground later. Our finding that after 10 and 11 years of exposure and programmatic learning, interference is the major persistent factor inhibiting acquisition, needs an explanation. We propose that a key difference between our research and much of the previous research lies in the languages examined. The close genetic and typological similarity of Hebrew and Arabic, according to the Crucial Similarity Measure (Wode, 1978), promotes transfer, both positive and negative, and this persists over the time range investigated.

In the domain of lexical acquisition, Haskel-Shaham et al. (2018) found a relatively high persistence of errors involving false cognates (lexical couplets similar in form but different in meaning) in the Hebrew interlanguage of Arabic-speaking college students. They call these false-cognate contexts—in which the L1 and L2 are expected to be similar but are not—'vulnerable states'. The researchers did not, however, go a step further to apply the term 'vulnerable state' to syntactic errors, although these were found to be the most persistent of the three domains they compared. We suggest that vulnerability of syntactic issues should be considered as a reason for their high persistence. This conforms with the Crucial Similarity Measure, according to which the similarity relations between languages, Hebrew and Arabic in our case, could be conducive to interference. We suggest that close similarity also causes persistence, in those domains where similarity is confusing.

6. LIMITATIONS

Our results show a surprisingly persisting effect of interference in the acquisition of syntax, which we suggested can be attributed to the typological similarity between Hebrew and Arabic being conducive to transfer. However, some limitations and caveats need to be considered. First, another way of explaining the results is that at the proficiency range level which we focus on, interference management is still not significant. This seems to include both time-points, whereas developmental errors decrease a little over this time range, especially 'haphazard' mistakes that are due to confusion and general lack of competence, rather than L2 irregularities. Further limitations concern our study itself: since our corpus consists of essays from only two time-points, separated only by a year, they may be indicative of only a particular part of the whole picture. It may be the case that over a longer time frame, we would have seen a decrease in the relative portion of interference-related syntactic errors. In order to make a stronger claim about the degree to which similarity between the L1 and L2 contribute to the persistence of interference, many more language pairs would need to be examined at progressive stages of instruction. Nevertheless, since these students had been studying Hebrew for over a decade, we believe that the results are noteworthy.

An additional possible limitation arises from our definition of interference errors, as any case in which an erroneous structure matches an Arabic structure. In other words—although everything that we coded as interference errors *can* be explained by interference from Arabic, it is not necessarily the only reason in each case, and we cannot rule out the option that some of the errors that we coded as 'interference' actually arise from other reasons, and only happened to have a chance similarity to a grammatical Arabic structure. Nevertheless, as explained in the discussion, since the group of errors that we labeled as having interference does indeed persist more than the others, we believe the most parsimonious explanation is that at least the majority of them do indeed contain interference, which contributed to their persistence.

Finally, our results admittedly depend on our own definition of error. As stated earlier, we distinguished errors from preferences, but as the line between the two is not always so clear-cut—and relies on attested variation among L1 speakers—it is possible that some of our exclusions would be considered errors by other researchers. More generally, we admit that a limitation of our study is our use of an abstract 'L1 Hebrew' as a benchmark, deviations from which were counted as errors or non-native like preferences. We are cognizant of the current critique in SLA of using native target-language norms as the baseline, anything short of which would be considered a failure (see Ortega, 2014). Selinker (2014, p. 228) and Granger (2015, pp. 13–16), for example, argue that SLA standards could be non-native in many cases. Indeed, in the case of English as a Lingua Franca, when most of the discourse occurs between non-natives, and the multitude of World Englishes is overwhelming, native proficiency is not essential, and 'expert' English is actually a more useful target (Granger, 2015, p. 16). In our case, however, the target language Hebrew is extremely limited in geographical scope and is relatively homogeneous. Unlike a global language such as English, Hebrew does not have a variety of endocentric norms and the degree of variation is far more limited. In terms of syntax, native Hebrew and expert Hebrew are practically indistinguishable.

Moreover, the community of Arabic speakers acquiring Hebrew has a very clear and practical notion of proficiency in Hebrew. The declared aim of educators, parents, and students alike, in the Arab community, is not simply functional competence for everyday communication, but access to higher education at the university, where academic Hebrew, as written by adult native speakers, is the baseline for assessment. In the absence of another operationalizable criterion for proficiency, we rely on this admittedly imperfect criterion as the baseline for our assessments.

7. CONLUSIONS AND FUTURE DIRECTIONS

The precise role which interference plays in SLA remains an open question. Our results add novel data to this debate, showing a high degree of syntactic interference and its persistence among students who had been studying Hebrew for a decade. We have suggested that these surprising results are due to the fact that unlike much

of the extant research, our study looks at two languages that are closely related, yet exhibit considerable differences in their syntax. In other words—close enough to be conducive for transfer, but different enough for it to cause persistence, in line with the Crucial Similarity Measure. Of course, further research is required in order to justify this proposal more fully. Since much of the field of SLA is dominated by research of English, our results highlight the need for SLA to investigate a broader range of languages, both as L1s and as L2s.

From a pedagogical point of view, a major implication of our findings is that once we recognize the link between interference and increased persistence, more attention can be attributed to these errors in teaching Hebrew to Arabic speakers. Depending on the preferred approach to instruction, these findings can be applied in appropriate methodologies. Taking the approach that we should target the learners' declared aims, rather than any preconceived ideal native speaker standard, we consider the aims of these students, which is generally higher education. As written language is assessed in Israeli universities according to norms of Standard written Hebrew, standard syntax is required. This may require some Focus on Form methods, in keeping with the recognition that "form-focused instruction might be relevant if that is what particular language learners want to be able to do in their language" (Ellis, 2019, p. 53). The syntactic domains most prone to persistent interference can be highlighted. This fits in with recommendations that "making learners aware of cross-linguistic differences will help with certain difficulties in the target language" (Odlin, 2003, p. 478) and "the closer the relation (between the languages), the more teaching can concentrate on the actual differences that exist" (Ringbom & Jarvis, 2009, p. 115). In our case, if the students are explicitly informed of the difference between Arabic and Hebrew in the specific areas where their mistakes are most persistent, they should be more aware of the problems and better equipped to avoid them in future usage. So, for example, prepositions may be taught along with the governing verbs, as already proposed by Haskel-Shaham et al. (2018, p. 21).

In communicative content-based approaches, 'vulnerable' discourse situations that are prone to communicative breakdowns due to L1-L2 differences could be tackled. Given the recent findings that discourse improves more than grammar and lexicon in the Hebrew writing of Arabic-speaking college students (Haskel-Shaham et al., 2018), it makes sense to exploit this relative advantage of discourse. In both approaches, focusing on the most common and persistent syntactic errors in both instruction and assessment should prove beneficial.

To conclude, both theory and instruction would clearly benefit from more research into the link between syntactic interference, developmental errors, and persistence. Our contribution to this relation in Arabic (L1) and Hebrew (L2) could well initiate research into this link in other dyads of similar languages, as well as dissimilar dyads. Future research could tackle larger samples, involving more types of errors, collected at different proficiency levels. Longitudinal studies ranging over a longer time-span than covered here would be especially enlightening for delving into these

issues, especially in a minority-majority sociolinguistic context, with SLA at high school targeting academic studies in the acquired majority language.

REFERENCES

AbiSamra, N. (2003). *An analysis of errors in Arabic speakers' English writings. Investigating writing problems among Palestinian students studying English as a foreign language* [Unpublished doctoral dissertation]. American University of Beirut.

Al-Khresheh, M. (2011). An investigation of interlingual interference in the use of 'and' as a syntactic coordinating structure by Jordanian EFL learners. *European Journal of Social Sciences, 18*(3), 426–433.

Bassal, I. (2007). Teaching Hebrew as L2 to Arab pupils: Interference from L1 and contribution of contrastive analysis. In S. Iliyan, Y. Avishur, N. Alkasem, & M. Hugerat (Eds.), *Jubilee volume in honor to Dr. Najeeb Nabwani: Studies in philosophy, culture & education* (pp. 101–133). Haifa, Israel: The Institute for Multi-Cultural Research and the Arab Academic College for Education in Israel [Hebrew].

Bataineh, R. F. (2005). Jordanian undergraduate EFL students' errors in the use of the indefinite. *The Asian EFL Journal Quarterly, 7*(1), 56–76.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing, 14*(3), 191–205. https://doi.org/10.1016/j.jslw.2005.08.001

Bohnacker, U., & Rosén, C. (2007). How to start a V2 declarative clause: Transfer of syntax vs. information structure in L2 German. *Nordlyd, 34*(3), 29–56.

Diab, N. (1996). The transfer of Arabic in the English writings of Lebanese students. *The ESP, Sao Paulo, 18*(1), 71–83.

Dulay, H. C., & Burt, M. K. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly, 8*(2), 129–136.

Ellis, N. C. (2019). Essentials of a theory of language cognition. *The Modern Language Journal*, *103*, 39–60. https://doi.org/10.1111/modl.12532

Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2021). A short history of SLA: Where have we come from and where are we going? *Language Teaching, 54*(2), 190–205. https://doi.org/10.1017/S0261444820000038

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing, 8*(1), 1–11. https://doi.org/10.1016/S1060-3743(99)80110-6

Gilquin, G. (2001). The integrated contrastive model: Spicing up your data. *Languages in Contrast, 3*(1), 95–123. https://doi.org/10.1075/lic.3.1.05gil

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research, 1*(1), 7–24. https://doi.org/10.1075/ijlcr.1.1.01gra

Haskel-Shaham, I., Klaus, R., & Tamir, R. (2018). Discourse—yes, grammar— no. Influence of Arabic mother tongue on Arab students writing in Hebrew. *L1-Educational Studies in Language and Literature, 18*, 1–26. https://doi.org/10.17239/ L1ESLL-2018.18.01.07

Henkin, R. (2010). *Negev Arabic: Dialectal, sociolinguistic and stylistic variation*. Otto Harrassowitz, Wiesbaden.

Henkin, R. (2020). Persistence of interference from L1 Arabic in written Hebrew. *L1-Educational Studies in Language and Literature, 20*, 1–31. https://doi.org/10.17239/L1ESLL-2020.20.01.15

Henkin-Roitfarb, R. (2011). Hebrew and Arabic in asymmetric contact in Israel. *Lodz Papers in Pragmatics, 7*(1), 61–100. https://doi.org/10.2478/v10016-011-0004-7

Heydari, P., & Bagheri, M. S. (2012). Error analysis: Sources of L2 learners' errors. *Theory & Practice in Language Studies, 2*(8), 1583–1589. https://doi.org/10.4304/tpls.2.8.1583–1589

Hinkel, E. (2005). Analyses of second language text and what can be learned from them. In E. Hinkel (Ed.), *Handbook of research in second language learning and teaching* (pp. 615–628). Mahwah, NJ: Lawrence Erlbaum Associates.

Jarvis, S. (2015). The scope of transfer research. In L. Yu & T. Odlin (Eds.), *New perspectives on transfer in second language learning* (pp. 17–47). Bristol, UK: Multilingual Matters.

Kalantari, R., & Gholami, J. (2017). Lexical complexity development from dynamic systems theory perspective: Lexical density, diversity, and sophistication. *International Journal of Instruction, 10*(4), 1–18.
Kaplan, R. B. (1966). Cultural thought patterns in intercultural education. *Language Learning, 16*(1), 1–20.

Long, M. (2003). Stabilization and fossilization in interlanguage development. In C. J. Doughty, & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 487–535). Oxford, UK: Blackwell. https://doi.org/10.1002/9780470756492.ch16
MacWhinney, B. (1992). Transfer and competition in second language learning. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 371–390). Amsterdam: North Holland.
MacWhinney, B. (2005). A unified model of language development. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford: Oxford University Press.
Noor, H. (1996). *English syntactic errors by Arabic speaking learners: Reviewed*. ERIC Document Reproduction service No. ED423660.
Odlin, T. (2003). Cross-linguistic influence. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 436–486). Oxford, UK: Blackwell.
Odlin, T. (2005). Crosslinguistic influence and conceptual transfer: What are the concepts? *Annual Review of Applied Linguistics, 25*, 3–25. https://doi.org/10.1017/S0267190505000012
Ortega, L. (2014). Ways forward for a bi/multilingual turn in SLA. In S. May (Ed.), *The Multilingual Turn: Implications for SLA, TESOL and Bilingual Education* (pp. 32-53). New York: Routledge.
Prior, A., Degani, T., Awawdy, S., Yassin, R., & Korem, N. (2017). Is susceptibility to cross-language interference domain specific? *Cognition, 165*, 10–25. https://doi.org/10.1016/j.cognition.2017.04.006
Ringbom, H., & Jarvis, S. (2009). The importance of cross-linguistic similarity in foreign language learning. In C. Doughty & M. Long (Eds.), *The handbook of language teaching* (pp. 106–118). Madden, MA: Wiley-Blackwell.
Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition, 30*(3), 333–357. https://doi.org/10.1017/S0272263108080480
Selinker, L. (2014). Interlanguage 40 years on: Three themes from here. In Z. H. Han & E. Tarone (Eds.), *Interlanguage : Forty years later* (pp. 221–246). John Benjamins Publishing Company.
Shawarbah, M. (2012). *A grammar of Negev Arabic: Comparative studies, texts and glossary in the Bedouin dialect of the Azazmih Tribe*. Semitica Viva, Otto Harrasowitz, Wiesbaden.
Shehadeh, H. (1998). The Hebrew of the Arabs in Israel. *Lěšonénu La'am, 49*(4), 168–180. [Hebrew].
Weinreich, U. (1968). *Languages in contact*. The Hague: Mouton.
Wode, H. (1978). Developmental sequences in naturalistic L2 acquisition. *Working Papers on Bilingualism, 11*, 1–31. (Reprinted in Hateh, E. (Ed.). 1978. 101–117.)
Zheng, Y. (2016). The complex, dynamic development of L2 lexical use: A longitudinal study on Chinese learners of English. *System, 56*, 40–53. https://doi.org/10.1016/j.system.2015.11.007