

# Context of learning and overcoming protracted instability at the interface in advanced L2 learning: Evidence from definite plurals in L2 Arabic

Arab Journal of Applied Linguistics

e-ISSN 2490-4198

Vol. 07, No. 02, 2022, 35-95

© AJAL

<http://www.arjals.com>

*Mahmoud Azaz<sup>1</sup>, University of Arizona, USA*

## Abstract

Interface properties have been found to exhibit protracted instability in advanced bilingualism and L2 learning (Sorace, 2000, 2005; Sorace & Filiaci, 2006). This study examines the role of learning context in modulating this instability in an interface property, namely the interpretation of definite plurals in L2 Arabic. Generic readings in English preverbal positions are expressed with bare plurals, which are ungrammatical in Arabic. Performances of two advanced English-speaking learners of Arabic in two learning contexts are compared: formal language instruction in an at-home (AH) setting and an extended study abroad (SA) setting. Results of two elicited production tasks reveal that whereas the advanced-AH group fluctuated between bare plurals (arguably transferred from L1) and definite plurals, the advanced-SA group opted for definite plurals. These results confirm Sorace's (2011) claim that protracted instability is real in interface properties in advanced bilingualism. They further suggest that input conditions and learning context play a crucial role in overcoming this instability. Active contact with the L2, which is characteristic of extended SA settings, is suggested to stabilize L2 forms and preempt competing L1 forms in interface properties.

**Keywords:** *Interface properties; Protracted instability; L1 transfer; Learning context; Generic plural nouns; L2 Arabic*

---

<sup>1</sup> Corresponding author: Mahmoud Azaz  
Email: [mazaz@arizona.edu](mailto:mazaz@arizona.edu)

## **Introduction**

Sorace and colleagues introduced the Interface Hypothesis (IH) to explain why certain constructions continue to present protracted instability and residual optionality in advanced bilingualism (Sorace, 2005; Sorace & Filiaci, 2006). These terms refer to the co-existence of two forms even though the target language (L2) permits only one. The ungrammatical form is arguably transferred from the first language (L1) (see Sorace, 2000, 2005; Valenzuela, 2006 for an overview). Constructions that require the integration of syntactic knowledge with knowledge from other domains, such as lexical semantics, discourse, and pragmatics, are characterized as interface properties. Two types of interfaces are differentiated: (i) internal interfaces, in which syntax interfaces with other internal modules in language (e.g., syntax-semantics interface), and (ii) external interfaces, in which syntax interfaces with other “higher” modules in the cognitive system (e.g., syntax-pragmatics/discourse interface) (see Sorace, 2011; White, 2011 for an overview). Phenomena that pertain to the interfaces are generally unstable and harder to acquire than phenomena that involve formal properties of the language system alone (Lozano, 2006; Sorace, 2011).

Experimental research has examined the role of L1 effects in the acquisition of interface properties. These effects have been found in syntax-pragmatics interface properties, such as subject null and overt pronoun distribution (Rothman, 2009), pronominal subjects (Haznedar, 2010), topicality and clitic doubling (Ivanov, 2009), and

subject-anaphoric reference scrambling (Hopp, 2009). Moreover, they have been found in syntax-semantics interface phenomena, such as aspectual interpretation (Gabriele, 2009; Montrul & Slabakova, 2003), quantifier scope (Dekydtspotter & Sprouse 2001), double-que questions (Cuza & Frank, 2011), container-content relations (Azaz & Frank, 2017), and article systems (Azaz, 2014, 2016; Cuza, Guijarro-Fuentes, Pires, & Rothman, 2012; Ionin, Ko, & Wexler, 2004; Ionin & Montrul, 2010; Ionin, Montrul, Kim, & Philippov, 2011; Sorace & Serratrice, 2009). In terms of article systems as a syntax-semantics interface properties, existent literature has primarily examined L1 effects on generic and specific interpretations of definite and indefinite/bare singular and plural nouns.

Recently, the focus on the L2 acquisition of interface properties has extended to exploring the conditions that modulate the protracted instability of target forms, especially in scenarios where L1 offers competing options (Sorace, 2000, 2005, 2011; Sorace & Serratrice, 2009; Valenzuela, 2006). The stability of the target forms is taken to mark the preemption of the L1 option. Consequently, retreat from persistent L1 effects is attained. Discussions have been complicated by two factors. The first is the inherent characteristics of the properties in question, whether purely syntactic, pertaining to internal interfaces, or pertaining to external interfaces. The second is the role of input properties (quantity and quality) that contribute to stabilizing the target forms in the interface properties. As a consequence, non-target forms start to shrink. Particularly relevant to input properties is the role of L2 learning contexts. Two learning contexts are

often distinguished: (i) typical formal language instruction in at-home (AH) settings, in which learners are exposed to language in addition to other areas of study; and (ii) the study abroad (SA) context that combines formal instruction and ample opportunities for continuous and active interaction with the native speech community (see Collentine & Freed, 2004; Freed, Segalowitz, & Dewey, 2004 for an overview).

The differential effects of these two learning contexts in linguistic development in general have been controversial (see DeKeyser, 1991; Huebner, 1995; Lapkin, Hart, & Swain, 1995, among many others). In part, this is due to the dearth of systematic comparative studies, both cross-sectional and longitudinal, that examine factors determining the superiority of each of these learning contexts. Additionally, studies that examine the effectiveness of the SA experience hypothesize that the solid lexical and grammatical base that advanced learners attain before their SA experience forms a threshold that facilitates the L2 acquisition of grammar subtleties (such as interface properties) without exhausting too many attentional resources (Isabelli & Nishida, 2005; Marqués-Pascual, 2011; Masuda, 2011; Segalowitz & Freed, 2004, among others). Lafford (2006) calls for putting the threshold hypothesis to test in comparative learning contexts. To date, very few studies make systematic comparisons between the differential effects of learning contexts on the acquisition of interface properties.

The present study contributes to this line of research by examining the role of learning context (AH versus SA) in modulating the protracted instability of target forms

in a syntax-semantics interface property. It scrutinizes the patterns of fluctuation between non-target forms (transferred from the L1) and target forms in the L2. Patterns of instability of the target forms are taken as a manifestation of competition with L1 forms. In comparison, the patterns of stability of the target forms are taken as a manifestation of the suppression of the L1 forms. Retreat from L1 effects is accomplished as a result of this suppression. The present study considers one particular case of the Arabic article system, which is the semantics of plural nouns. Existent literature characterizes definite and bare plurals cross-linguistically as pertaining to the syntax-semantics interface because the shift between bare and definite plurals is associated with a subtle shift in the interpretative possibilities, including generic, specific, and existential readings (see Sorace & Serratrice, 2009; Cuza et al. 2012). This shift is further complicated by the intricacies of word order whether preverbal or postverbal (see Cuza et al. 2012; de Garavito, 2013).

Specifically, the present study asks whether advanced learners who studied Arabic exclusively in a formal classroom setting at home would show a significantly different pattern of protracted instability of the target forms (definite plurals that encode a generic meaning) when compared to their advanced counterparts of comparable proficiency who studied Arabic abroad in an extended immersion setting. In addition, it examines whether the participants' patterns of in/stability are associated with their language contact profiles in their respective learning contexts. The role of learning

context should deepen our understanding of the linguistic and non-linguistic factors that modulate cross-linguistic influence effects at the interfaces.

The article is structured as follows: Section (2) provides relevant background in four areas: (i) AH and SA learning contexts in L2 studies, (ii) L2 acquisition of interface properties, (iii) generic and specific plurals in English and Arabic, and (iv) article semantics in L2 studies. Section (3) pertains to the study itself, including research questions, participant profiles, learning contexts, and materials. Section (4) provides group results and individual analysis. Section (5) offers the discussion, which considers the implications of the results to the acquisition of interface properties. Section (6) provides a brief conclusion along with implications for future research.

## **Background**

### *AH and SA learning contexts in L2 studies*

The effects of learning contexts have a long history in L2 learning theories. Experimental studies have explored aspects of linguistic development in three primary learning contexts: (i) formal language instruction in at-home (AH) institutions, (ii) intensive domestic immersion (IM) at home mostly in summer programs, and (iii) the study abroad (SA) setting (see Collentine & Freed, 2004 for an overview). A detailed discussion of the differential effects of these contexts is beyond the scope of this article. However, a general conclusion that emerges from earlier studies (Carroll, 1967; Freed, 1995) as well as from

more recent studies (Freed, Segalowitz & Dewey, 2004) is that the SA experience is typically superior to that of AH in terms of developing oral fluency and overall proficiency of the target language.

Recently, a large body of research has explored the role of the SA experience on specific aspects of linguistic development. These studies conclude that SA experiences do not benefit all aspects of linguistic development in the same way. Evidence was found in support of the impact of the SA context on oral fluency (Freed, Segalowitz, & Dewey, 2004), L2 pronunciation (Llanes, Mora, & Serrano, 2016), and interactional competence (Masuda, 2011). In the domain of grammatical properties, results of comparative studies are far from conclusive. For example, a large number of studies have shown that the SA context is not superior to the AH context in making overall gains (see Collentine, 2004; DeKeyser, 1991; Lafford, 2006; Segalowitz, Freed, Collentine, Lafford, Lazar, & Díaz-Campos, 2004 for a review). However, few studies show an advantage for the SA context in terms of grammatical abilities. Isabelli and Nishida (2005) found superior control of the subjunctive by English-speaking learners of Spanish in the SA context, but not in the AH context. Very few studies sought to examine the gains in grammatical properties that vary in terms of their inherent characteristics. Marqués-Pascual (2011) is one of these few studies. It found that whereas both AH and SA advanced learners of Spanish make similar gains in verb morphology, subject omission (a property pertaining to the syntax-pragmatics interface) benefits more from the SA experience.

### *The L2 acquisition of interface properties*

The Interface Hypothesis (IH) was put forward by Sorace and colleagues to explain why certain constructions continue to present residual optionality and protracted instability at the advanced stages of L2 learning (Sorace, 2005, 2006, 2011; Sorace & Filiaci, 2006). These constructions are often less likely to be acquired completely. According to White (2011), there are two versions of the IH. In version I, which is the original proposal by Sorace and Filiaci (2006), narrow structures are “more privileged” than interface constructions in which syntax interfaces with other internal modules. In version II, a fine-grained distinction is made between two types of interfaces: internal and external. Internal interface constructions refer to those in which syntax interfaces with other internal modules in language (e.g., morphology, phonology, and semantics). External interface constructions refer to those in which syntax interfaces with other “higher” modules in the cognitive system (e.g., pragmatics and discourse).

The syntax-pragmatics/discourse interface (external interface) and the syntax-semantics interface (internal interface) have caught increasing attention in L2 studies. Syntax-semantics interface phenomena are thought to provide a less significant source of protracted instability than syntax-pragmatics interface phenomena. However, research results regarding each of these two interfaces are mixed. In the syntax-pragmatics interface, some studies provide evidence that null and overt subject distribution is vulnerable (Belletti, Bennati, & Sorace, 2007; Belletti & Leonini, 2004; Hertel, 2003;

Lozano, 2006; Sorace & Filiaci, 2006). However, Rothman (2009) examines the same phenomenon and concludes that it is indeed acquirable by advanced learners. Experimental work on other syntax-pragmatics interface phenomena, such as clitic doubling, continues to provide mixed findings. For example, Tsimpli and Sorace (2006) report protracted instability in the use of clitic doubling by advanced Russian-speaking learners of Greek. Ivanov (2009) and Slabakova and Ivanov (2011) conclude that advanced English speakers acquire the discourse constraints on topicality and clitic doubling in Bulgarian. Valenzuela (2005, 2006) shows that advanced-English speaking learners of Spanish do not acquire discourse constraints of non-specific topics with clitic.

Regarding the acquisition of syntax-semantics interface phenomena, studies have provided mixed findings. The characterization of the syntax-semantics interface in L2 acquisition owes much to the work of Dekydtspotter and colleagues. They have examined the interpretative complexities of word order variants in L2 French by English-speaking learners in a series of studies. These included process-result distinction in dyadic nominals (Dekydtspotter, Sprouse, & Anderson, 1997); combined *'how many'* extractions (Dekydtspotter, Sprouse, & Swanson, 2001), and quantification at distance (Dekydtspotter & Sprouse, 2001). The general conclusion that emerges from researching this set of constructions is that advanced learners are able to acquire the interpretative possibilities associated with word order variants. Studies conducted on the acquisition of other syntax-semantics interface properties also yield mixed results. For example, Azaz

and Frank (2017) conclude that the English container-content distinction (*coffee cup* versus *cup of coffee*) poses protracted instability and persistent difficulties for Spanish-speaking and Arabic-speaking learners. Borgonovo, De Garavito, & Prévost (2008) conclude that the acquisition of specific vs. unspecific interpretative possibilities of mood (indicative vs. subjective) in Spanish relative clauses is quite problematic for French speaking learners.

### ***Definite and bare plurals in Arabic and English: A syntax-semantics interface property***

Modern standard Arabic (MSA) and English are classified as article languages. In Arabic, the definite article *al-* is a proclitic that is attached to the noun it defines. In English, the definite article *the* is an independent word. Comparative work on the syntax-semantics interface of definiteness in MSA and English has shown clear differences in how definite and bare nominals are interpreted in preverbal and postverbal positions (see Fehri 2004 for an overview). Generally speaking, three interpretative possibilities are discussed in relation to article semantics: generic, specific, and existential.

For genericity or kind reference, English uses a wide range of noun phrases in the preverbal position, one of which is bare plurals. More importantly, a generic reading is not available for definite plurals as illustrated in (1).

- (1) a. *A dog* barks.  
 b. *The dog* barks.  
 c. *Dogs* bark.  
 d. *#The dogs* bark. (Generic)

Since the early work of Smith (1975), the semantic and the pragmatic distributional differences between these three nominals has been controversial in linguistic analysis studies (see Chierchia, 1998; Dayal, 2004; Greenberg, 2007; Krifka, Pelletier, Carlson, ter Meulen, Chierchia, & Link, 1995). Overall, the choice between these three variants is determined by what each nominal predicates. Two predication possibilities have been discussed in the literature. The first involves predicates that apply to the class/type as one entity or unit. These include predicates such as “*be extinct*,” “*be widespread*,” and “*be common*.”

- |     |                                    |   |
|-----|------------------------------------|---|
| (2) | <i>A squid</i> lives on seaweed.   | [indefinite singular NP: kind-referring]  |
|     | <i>The squid</i> lives on seaweed. | [definite singular NP: kind-referring]    |
|     | <i>Squids</i> live on seaweed.     | [bare plural NP: kind-referring]          |
| (3) | * <i>A dodo</i> bird is extinct.   | [indefinite singular NP: #kind-referring] |
|     | <i>The dodo</i> bird is extinct.   | [definite singular NP: kind-referring]    |
|     | <i>Dodo birds</i> are extinct.     | [bare plural NP: kind-referring]          |

For these predicates, definite singulars and bare plural subjects are grammatical, but indefinite singular subjects are not (see Krifka et al., 1995). The second possibility concerns predicates that are not necessarily characteristic of the type/class as a unit or

entity. In recent work on the acquisition of genericity distinctions, Ionin, Montrul, and Crivos (2013) differentiate between genericity at the NP level and genericity at the sentence level to further clarify this distinction. These are exemplified in (2) and (3) from Smith (1975).

The use of the definite article with plural nouns in English in preverbal positions yields a specific reference/maximality as illustrated in (4).

- |     |                       |  |
|-----|-----------------------|--|
| (4) | <i>Dogs</i> bark.     | [Indefinite plural NP: kind-referring]   |
|     | <i>The dogs</i> bark. | [Definite plural NP: specific reference] |

In the semantics literature, Chierchia (1998) and Dayal (2004) examine how kind-denotation and maximality are expressed in English and Romance languages. According to them, the definite marker in English plurals lexicalizes specificity/maximality, the semantic operation that maps specific individuals to the set in the category. For example, “*the lions*” denotes a group of lions in a certain discourse. In other Romance languages such as Spanish, the use of definite plurals has a dual function, as it denotes kind reference in addition to specific reference.

In MSA (and also in Arabic dialects), for generic reference to types, definite singular and plural nouns are interchangeably used in the preverbal position (Fehri 2004, p. 44) as illustrated in (5) and (6).

- (5) **al-kalb-u**                      yanbaḥ-u  
**def-dog-nom.**                      barks  
 “The dog barks.”
- (6) **al-kilaab-u**                      tanbaḥ-u  
**def-dogs-nom.**                      bark  
 “Dogs bark.”

It has been argued by semanticists (see Longobardi, 2001) that Arabic definite plurals behave similarly to those in Romance languages. They refer to generic and specific interpretations. Like Spanish, for instance, Arabic uses definite plurals for maximal denotation when specific reference is made to a certain group, members, exemplars, or entities. This is illustrated in (7).

- (7) **al-kilaabu**    ʔallatii    raʔajtuhaa    ʔams    fii    bayt-i    ʕammii    lam    tanbaḥ  
**def-dogs**    which    I saw    yesterday    in    house    uncle    not    bark  
 “The dogs which I saw yesterday at my uncle’s house did not bark.”

Relating this typological difference between MSA and English to semantic accounts in the literature, definite plurals (e.g., *the lions*) in English lexicalize maximality, but they do not encode generic readings or kind reference. In MSA, definite plural nouns encode maximality/specificity and genericity.

Interestingly, in MSA bare plurals are grammatical in the preverbal position, but they are assigned an existential reading as illustrated in (8).

- (8) kilaab-un                      kabiira-tun                      tʿaaradat-nii                      ʔams  
 dogs.indef-nom.                      big                      chased-me                      yesterday  
 “Big dogs chased me yesterday.”

This example shows clear differences between MSA and English in terms of how bare and definite plurals are interpreted. In line with existent L2 acquisition studies, (e.g., Cuza et al., 2012; Sorace & Serratrice, 2009), the properties of bare and definite plural nouns in Arabic and English pertain to the syntax-semantics interface. This is because the choice between bare and definite plurals is associated with a subtle interpretative shift (generic, existential, or specific) that is further complicated by word order constraints.

*Article semantics in L2 acquisition studies*

Studies conducted on the acquisition of article semantics have focused mostly on the role of L1 effects in the interpretation of definite and bare plural nouns, whether generic or specific. Research in this area owes much to the work of Ionin and colleagues on multiple language combinations. In one study, Ionin, Montrul, and Crivos (2013) used a bi-directional framework with English-Spanish and Spanish-English learners to explore the effects of L1, both English and Spanish, on the interpretation of bare and definite plurals in L2. Two tasks were used: a meaning-based task (that evaluated learners’ interpretations of definite plurals) and a form-focused task (that examined learners’

judgments of the acceptability of definite and bare plurals in generic and specific contexts). The study reports L1 effects in both directions at lower proficiency levels. At higher proficiency levels, recovery from these effects was possible. Three sources were speculated for this recovery: input and instruction, the participants' underlying sensitivity to the semantics of the specific versus generic reference, and general cognitive mechanisms.

In another study, Ionin and Montrul (2010) ask whether Spanish-speaking learners of English would over-accept the generic readings of English definite plurals to a greater extent than Korean-speaking learners of English. The comparison with Korean is important because it does not have a definite article. Using an acceptability judgment task and a truth value judgment task, the study offers evidence for L1 effects in the interpretation of generic plurals: Spanish-speaking learners of English tend to assign generic readings to English definite plurals more than their Korean-speaking counterparts. When more advanced learners (11 Spanish-speaking and 9 Korean-speaking) who were immersed in an English-speaking environment for a short length of residence were tested, both groups demonstrated different patterns of recovery, as only half the participants in the two groups performed in a target-like manner. The study concludes that immersion is no guarantee of target-like performance. Similar L1 effects were found by Montrul and Ionin (2010). Based on four written tasks (acceptability judgment, truth-value judgment, picture sentence matching, and sentence-picture

acceptability judgment), results showed that English-speaking learners of Spanish (L2 proficiency rating was 3.8/5.0 and L1 proficiency rating was 5.0/5.0) demonstrated evidence of dominant L1 transfer in their interpretation of plural nouns.

Cuza et al. (2012) examined the extent to which advanced English-speaking learners acquire Spanish definite plurals with generic and specific readings (e.g., *los gatos* 'the cats' vs. *gatos* 'cats'). Starting from a learnability assumption that learning is easier than unlearning, the task of English-speaking learners of Spanish was characterized as a remapping task of expansion, since they come to permit generic interpretations with specific interpretations to definite plurals. In their L1 English, only the latter reading is possible. The acquisition of this expansion means the loss of generic interpretation to bare plurals in L2 Spanish. Using a context felicitousness task, results showed that English-speaking learners of Spanish demonstrated target knowledge of the semantics of definite and bare plurals. However, while there were similar statistical differences in performance on definite plurals, the advanced group did not disallow the generic reading to the bare plurals to the same degree as native speakers. These results were interpreted in light of feature resetting.

Serratrice, Sorace, Filiaci, and Baldo (2009) explored bilingual children's ability to give metalinguistic judgment on the grammaticality of plural nouns in generic and specific contexts in English and Italian. This study, although it focused on early bilingualism in children, is particularly relevant to the present study given its focus on

input conditions in the acquisition of article semantics. It examined, among other things, whether the typological relatedness of the bilinguals' two languages significantly affect performance. Using an acceptability judgment task with pictures, the English-Italian bilinguals living in the UK and in Italy and the Spanish-Italian bilingual children in Spain judged whether the sentence they had heard was grammatically correct or incorrect in English and Italian. The generic interpretations were contextualized using adjuncts such as '*in general,*' and the specific interpretations were contextualized using adjuncts such as '*here.*' Overall, the results demonstrate that all the groups performed more accurately in the specific context than in the generic context. For the generic condition, the study found strong evidence that knowledge of English affected the bilinguals' ability to discriminate between grammatical and ungrammatical sentences in Italian.

In comparison to the L2 studies on Romance languages, the acquisition of article semantics in Arabic is an under-researched area (see Alhawary, 2009 for an overview of Arabic SLA morphosyntax). Most of the existent studies were conducted using the contrastive analysis and error analysis approaches without exploring how certain semantic aspects of the article system are acquired. In an early study, Al-Ani (1972-1973) analyzed a limited number of written compositions in search for major errors. Three types of errors were identified: orthographic and phonological, diction and dictionary usage, and grammatical errors. Errors in the definite marker in noun phrases, although listed as one of the major errors, were attributed to three sources: interference, overgeneralization

strategies, and competence versus performance errors. The study does not identify clear reasons for the definite marker errors by English-speaking adult learners. It concludes that it was difficult to identify the categories of these errors and their true sources. Rammuny (1976), in a more extensive study than that of Al-Ani, analyzes common errors in written Arabic compositions in proficiency tests. He identifies three categories of errors: phonological, lexical, structural, and stylistic. Errors in definiteness are categorized as structural errors along with 16 other error types under the same category. Along with these diverse error types, causes remain unclear, as they are attributed to L1 interference, competence and performance issues, and insufficient teaching-learning strategies.

Only one study (Azaz, 2016) specifically addresses the acquisition of certain semantic aspects of the Arabic definite article by English native speakers. The focus of this study is L1 effects on the acquisition of three categories of definite singular determiner phrases that encode generic reading in Arabic. These show (i) similarities (determiner phrases that denote unique entities such as *the sun*), (ii) differences (determiner phrases that denote abstract concepts such as *love*), and (iii) structural overlap (determiner phrases that denote kind with singular nouns, such as *the lion*). Using grammaticality judgment and forced choice tasks, the study reports L1 effects at low proficiency levels in terms of definiteness, although not full, for the types that show similarities and differences. For kind-denoting singulars, the study found that whereas

beginning English-speaking learners fluctuate between singular definite-marked nouns and singular bare nouns, advanced-speaking learners tend to use the target definite singulars.

### **The present study**

In summary, article semantics studies primarily examine L1 effects on the interpretation of bare and definite plurals (generic or specific) as a syntax-semantics interface property in Romance languages and English. The general picture that emerges from this literature is that the semantics of bare and definite plurals is a vulnerable domain in adult L2 acquisition due to crosslinguistic influence effects and the syntax-semantics interface complexities. Only Ionin and Montrul (2010) consider the role of immersion in retreat from L1 effects and conclude that it is no guarantee of target-like performance. More importantly, no studies have been conducted to examine the patterns of protracted instability of the target forms (definite or bare) and retreat from L1 effects in different learning contexts. Broadly speaking, the role of input (quantity and quality) has always been at the center of discussions of residual optionality and protracted instability in constructions that are sensitive to the interfaces (Sorace, 2011, 2014; Sorace & Serratrice, 2009; White, 2011). However, to date no studies have been conducted to scrutinize the role of the learning context in modulating patterns of protracted instability and residual optionality in the acquisition of interface phenomena.

The present study contributes to filling this gap. It examines the role of the learning context in modulating the patterns of protracted instability of generic definite plurals at advanced proficiency levels in L2 Arabic. It scrutinizes fluctuation between non-target bare plurals (arguably transferred from L1 English) and target definite plurals under two learning conditions: (i) at-home setting that focuses on formal classroom and teacher-learner interactions, and (ii) the study abroad setting that combines formal classroom instruction and continuous exposure, as well as active use of L2 in the native speech community.

***Research questions and predictions:***

This study attempts to answer the following questions:

- i. Do advanced learners of Arabic who studied in an at-home classroom setting (advanced-AH) exhibit a significantly different pattern of protracted instability of generic definite plurals from advanced learners who were immersed in a study abroad setting (advanced-SA)?
- ii. If so, which learning context is preferable in order to overcome patterns of protracted instability?

For the first research question, according to the original proposal IH, it is expected that the syntax-semantics interface complexities will result in persistent protracted instability of generic definite plurals for the advanced-AH group. Using definite plurals

in generic contexts in Arabic requires the integration of syntactic knowledge (the choice between bare and definite plurals) and semantic knowledge (generic and specific readings). The advanced-AH group is expected to fluctuate between definite and bare plurals when the generic meaning is elicited. The fact that the semantics of definite and bare plurals in Arabic is neither integrated in textbooks nor taught in L2 classes may exacerbate this protracted instability. Due to the limited nature of input (quality and quantity), the L1 option (bare plurals) is expected to compete persistently with the target form (definite plurals) in the at-home learning context. In comparison, the advanced-SA group is expected to display a marginal degree of instability. Continuous exposure, active use of the L2, and reduced use of L1 in the SA environment is expected to result in preempting the L1 competing option and overcoming the interface complexities. Subsequently, they are expected to demonstrate a target-like performance.

For the second research question, the language contact profiles in the two learning contexts will be further analyzed. It is expected that the variable patterns of in/stability of generic definite plurals by the two advanced groups will be associated with their variable language contact profiles (in class and out-of-class) in their respective learning contexts. In particular, the continuous exposure and active use of L2 and the reduced use of L1 by the advanced-SA group should be reflected in their out-of-class language contact profiles. The SA context is expected to be preferable to overcome patterns of protracted instability of the target-like forms.

### *Participant profiles*

A total of 30 individuals participated in this study. They were divided into three groups as follows: (i) an advanced-proficiency group who studied Arabic only in an at-home setting (advanced-AH; N = 10); (ii) an advanced-proficiency group who studied Arabic abroad (advanced-SA; N = 10); (iii) and a group of native Arabic speakers (N = 10). The two advanced L2 groups were native speakers of English. The 10 participants in the advanced-AH group were selected from an initial pool of 15 participants. Five participants were excluded because they did not establish comparable proficiency with the rest of their group. The AH participants were recruited from a larger pool of upper

**Table 1**  
Participants' Background Information

Group	N	Age at testing		Years of instruction		Self-ratings	
		Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Advanced-AH	10	23.25 (5.24)	19-36	3.00 (0.77)	3-3.5	3.25 (0.59)	3.00-4.00
Advanced-SA	10	26.81 (2.71)	24-34	3.54 (0.85)	3.5-4	3.75 (0.40)	3.50-4.50
Native	10	32.00 (2.50)	28-37	NA	NA	NA	NA

levels of Arabic in an Arabic program in a southwest university in the United States. The 10 participants in the advanced-SA group were selected from an initial pool of 12 participants. Two participants were excluded because they reported learning Arabic abroad for more than three years. In addition, they worked as tutors of Arabic as a foreign language. It was concluded that this would affect the group homogeneity.

The academic majors of the two L2 groups spanned humanities, social sciences, and natural sciences. They were studying Arabic as a foreign language. The majority of the advanced-AH group (8 out of 10) were undergraduate students, and 2 were graduate students, whereas the majority of the advanced-SA (8 out of 10) were graduate students, and two participants were undergraduate students. The native control participants were graduate students from the Middle East and North Africa in the institution where the study was conducted. Before the beginning of the study, the two advanced L2 groups completed a language history questionnaire asking about their first language, classroom levels, gender, age, history of foreign language learning, and the duration of their study abroad (if applicable). Those who reported learning Hebrew or Romance languages (especially Spanish) at the advanced level were excluded. This procedure was followed because Hebrew and Romance languages show striking similarities to Arabic in the behavior of definiteness in plural nouns. Information about the participants is provided in Table 1.

At the time of testing, the L2 advanced groups were placed in their respective proficiency levels as determined by self-ratings and the American Council on Teaching of Foreign Languages (ACTFL) oral proficiency scale (Swender, Conrad, & Vicars, 2012). This scale was selected for its integration of components that assess linguistic, communicative, pragmatic, and discoursal competence. Therefore, it provides a holistic measure of proficiency as a complex construct. All participants in the L2 advanced groups performed at the advanced level in individualized unofficial oral proficiency interviews. This was taken as a basis for establishing comparable proficiency in the two groups. Additionally, the participants' comparable proficiency scores were supplemented with their proficiency self-ratings. The participants were asked to self-rate their proficiency on a scale from 1.0 (low) to 5.0 (superior). Self-ratings are extensively used as a measure of proficiency, and they have been shown to correlate with linguistic performance (e.g., Flege, MacKay, & Piske, 2002). These ratings are provided in Table (1). Further details of the participants' history of Arabic learning in their respective learning contexts are provided in the next section.

### *Learning contexts*

#### *AH context*

At the time of testing, all the advanced-AH participants had completed a minimum of three years of Arabic, in which they covered the first, second, and a substantial portion

of the third part of *Al-Kitaab Book Series* (Brustad, Al-Batal, & Al-Tonsi, 2001), which is one of the most common textbooks for teaching Arabic as a foreign language in the United States. They did not study Arabic abroad. In total, they had an average of around 280 contact hours of instruction. In the Arabic program where the study was conducted, elementary (first year), intermediate (second year), and advanced (third year) learners are taught Arabic for five hours a week following an integrated approach which combines Standard Arabic and colloquial Arabic (one dialect) in the same course, but on separate days. Standard Arabic is taught for four hours a week, and colloquial Arabic is taught for one hour a week. The underlying philosophy of this integration seeks to provide Arabic learners with a communicative command that enables them to function in both varieties. Generally, instruction is conducted in Arabic, while English is kept to a minimum. The four language skills are integrated. Although the main teaching method can be described as communicative in essence, explicit grammatical explanations are provided as needed. According to many instructors in the program, these explanations are needed due to the highly inflected system of Arabic. In addition to formal classroom instruction, the program offers extracurricular activities. These include Arabic conversation circles and cultural club events (4-5 events) every semester, during which various aspects of Arab culture are presented.

### *SA context*

At the time of testing, all the advanced-SA participants had completed a minimum of a full semester or equivalent (approximately 14 weeks) in a very intensive immersion environment in an Arabic-speaking country. Before that, they had all studied Arabic in formal settings for three years in the United States where they all studied *Al-Kitaab Book Series*. Eligibility criteria for their study abroad program required at least three years of formal instruction of Arabic. Their SA experience was quite diverse and intensive. It rested on a pedagogical philosophy that developed the four language skills at the advanced level in an integrated fashion. They had 16-18 contact hours per week that were distributed as follows: (i) Modern Standard Arabic for 6 hours per week (mostly reading texts selected from the media), (ii) colloquial Arabic for 6 hours per week (mostly speaking), (iii) writing skills for 3 hours per week in a workshop format, and (iv) an obligatory office hour for 1.5 hours. Listening activities were integrated with speaking and reading. On the weekends, they were expected to spend 8-10 hours on homework assignments. In addition, their SA program included three important out-of-class components: (i) a lecture series in which they engaged in conversation and discussion with presenters in MSA, (ii) a cultural program in which they developed a short community-based project outside the classroom, and (iii) language partnership in which each learner was paired with a language partner from the local community to speak only in the target language.

### *Tasks and procedures*

Three instruments were used in the present study: two prompted production tasks (sentence completion and oral narrative) and a language contact profile (LCP) survey. Prompted elicitation tasks are generally designed to elicit particular language forms or structures while the participants' focus is on meaning (Gass & Mackey, 2007; Rebuschat & Mackey, 2013). These include elicited imitation, interactive tasks, role-plays, picture descriptions, and oral narratives. Unlike form-focused tasks, such as grammaticality judgment and forced choice that either tap into explicit knowledge or allow strategic guessing, elicited production tasks are thought to tap into implicit knowledge because participants focus on meaning while processing L2 forms. Highly frequent noun plurals in the two elicitation tasks were pooled from the first part of *Al-Kitaab fii Ta'allum al-'Arabiyya: A Textbook for Beginning Arabic* (Brustad et al., 2001). These noun plurals were distributed among regular and irregular, masculine and feminine, and human and non-human. Further details of these two tasks are provided below.

#### *Prompted sentence completion*

The prompted sentence completion task asked the participants to read each sentence to understand the meaning established and provide the single missing word with the help of a picture. Participants were required to provide their responses as spontaneously as possible in speaking and in writing in individualized interviews that were recorded.

Thirty target items were distributed into two conditions: definite plural nouns that encode generic readings ( $N = 15$ ), and definite plural nouns that encoded specific readings ( $N = 15$ ). To facilitate reading the sentences and to further clarify the context (whether generic or specific), translations of potentially unfamiliar vocabulary (other than the target items) were provided.

(9) a. kulluna naʕrifu ʔanna  yaaliban tasiiru bi-ʔalbitruul  
a

We all know that \_\_\_\_\_ often run on petroleum  
“We all know that **cars** often run on petroleum.”

b. ʔanaa darastu fii al-baraaziil al-sana al-maadʕija  
I studied in Brazil year last



\_\_\_\_\_ ʔallaðiina darastu maʕahum kaanuu ɖʒajjiidin ɖʒiddan  
\_\_\_\_\_ which I studied with them were good very  
“I studied in Brazil last year. **The students** with whom I studied were very good”.

In line with Serratrice et al. (2009), generic readings were contextualized using the following prefaces: ‘*We all know that...*’ and ‘*As we all know.*’ They were further contextualized by integrating adverbs that encode genericity such as ‘*generally,*’ ‘*mostly,*’ and ‘*always*’ in the preface sentences. All the target generic plurals were presented in the

preverbal position. Specific readings were contextualized by introducing opening preface sentences, which were followed by relative clauses that further specify the target plurals. Additional modifying phrases such as *'in this place'* and *'in this city'* were also used. The generic and specific conditions in the task were scrambled with fillers that tested for subject–verb agreement, noun–adjective agreement, demonstrative–noun agreement, and verb tenses. All the test items and the fillers were scrambled throughout the test. Two examples of the generic and specific readings are provided in (9a) and (9b), respectively.

In (9a), the preface *'We all know that...'* establishes a generic reading. In contrast, the sentence in (9b), offers a specific reading established by the preface sentence and the modifying relative clause. In both contexts, the participants were expected to provide definite plurals: *al-sajjaaraat* 'the cars' for (9a) and *al-t'ullaab* 'the students' for (9b). The average time taken to complete the task by the two advanced groups ranged between 15-30 minutes. Singular nouns, incomplete plurals, and missing data were excluded from the analysis. For the advanced at-home group, only 1.82% of the total number of target items were excluded in the generic condition. Only less than 1.00% were excluded in the specific condition. For the advanced SA group, less than 1.00% of the data was excluded in the generic condition, and no data items were reported missing for the specific condition.

### *Prompted oral narrative*

The prompted oral narrative task required the participants to respond orally to ten short written prompts. Their narratives were recorded in individualized testing sessions. Specifically, they were asked to provide generic statements about the target plural noun/s in each prompt and support their statements with elaboration in about 2-3 minutes. The methodological benefit of this elicitation technique is that it engages the participants in a meaning-based task while the accuracy of the target forms is evaluated. For each prompt, when participants produced at least 3 clear frequencies of the target generic noun plurals, the interviewer let them finish their sentence and moved on to the next prompt. Two examples are provided in (10a) and (10b) below:

(10) a. In general, do you agree that big cities are better than small cities? Why/why not?

b. In general, do you think kids/children love colors? Why/why not?

The expected responses should provide definite plurals in reaction to each prompt for first mention in the context and for further mention, if the generic reading is maintained. For each prompt, all the frequencies of the target noun plurals (bare or definite) that encoded generic meanings were analyzed. If a participant did not produce at least two frequencies of the target plural nouns, the interviewer asked 2-3 follow-up questions in English to

elicit frequencies of generic target noun plurals. The average length of the interviews ranged between 16-25 minutes.

### *Language contact profile (LCP) survey*

To characterize the L2 advanced groups' language contact profiles (LCPs), information about their in-class contact hours across the four language skills was collected (number of days per week and number of hours per day). This information was supplemented with information about their out-of-class contact hours in their respective learning contexts. The out-of-class LCP survey developed for Spanish by Freed, Dewey, Segalowitz, and Halter (2004) was adapted for the purpose of this study. The out-of-class LCP survey gathered information about the curricular and extracurricular activities, both interactive and non-interactive, in which the participants were engaged. Specific information about the frequency of use (number of days per week and number of hours per day) of Arabic across the four language skills was provided. At the end of this LCP, participants in each group were asked to add their comments, if any, on their learning experiences and how they met their expectations as advanced language learners of Arabic.

### *Scoring*

In scoring the two prompted production tasks, special focus was given to the pronunciation of the definite article *al-* in definite plurals. This is because in the oral mode, the second sound *-l* of the definite article varies as a result of assimilation in certain

phonological contexts. If a participant thought that the noun followed a regular pattern and provided the plural form (bare or definite), his/her response was scored. For example, the word *madiina* 'city' in MSA follows an irregular (i.e., broken) pattern, and it is pluralized as *mudun* 'cities.' If a participant provided a regular pattern (e.g., *madiinaat*) with the plural ending *-aat*, the response was still considered valid. Mistakes that did not have to do with definiteness and plural marking were tolerated. The distribution of bare and definite plurals was calculated. A target-like definite plural was given a value of 1.00 and a non-target-like bare plural was given a value of 0.0.

In scoring the LCP survey, every hour spent in class or out-of-class per week was taken as a unit of analysis. The average number of hours per participant and per the whole group was calculated. A value of 1.0 was given to every contact hour spent, and 0.5 was given to half an hour. A value of 0.0 was given when a participant reported no time.

## **Results**

### ***Research question 1: Contrasting patterns of protracted instability***

To answer the first research question about the patterns of in/stability of the target definite plurals and retreat from L1 effects, the averages of bare and definite plurals in the two tasks were calculated. Results are provided below.

### *Elicited sentence completion*

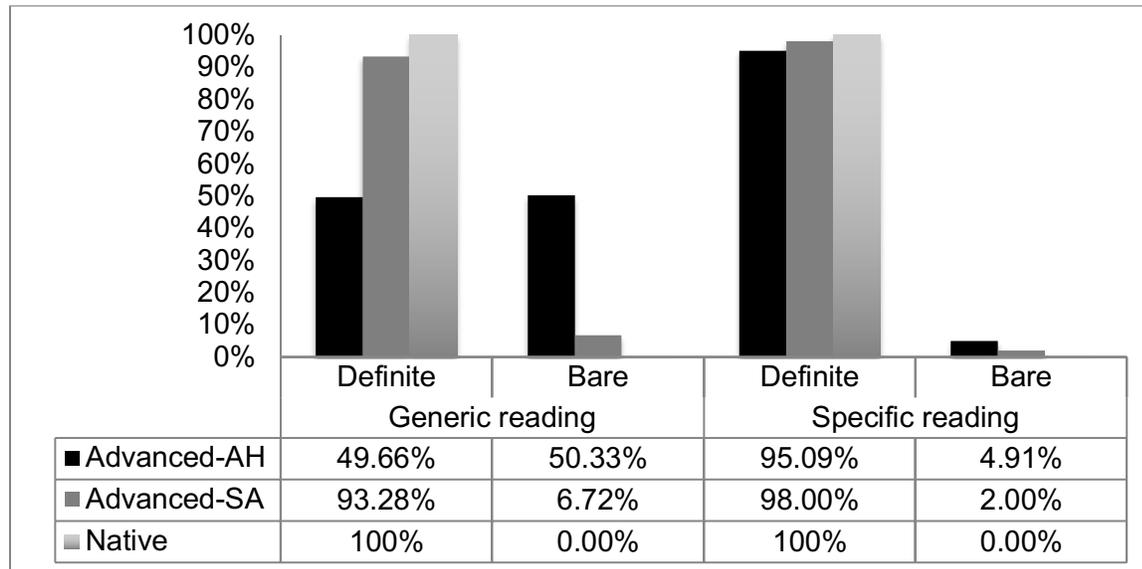
As illustrated in Figure 1, for the specific reading condition, the average scores of definite plurals were expectedly near ceiling for the two advanced groups. This target-like performance was expected, given the fact that in both Arabic and English definite plurals are used to encode specific readings. Results of the generic reading condition presented a different pattern. For the advanced-AH group, their performance overall was far from being target-like. The average score of target definite plurals was surprisingly at chance (49.66%; SD: 0.31), and the average score of non-target bare plurals was slightly higher (50.33%; SD: 0.31) with no difference between both averages in a paired-samples *t*-test:  $t(9) = 0.9692, p = 0.3608$ . This pattern of fluctuation between bare and definite plurals showed that the advanced-AH participants were exhibiting protracted instability as a manifestation of persistent L1 effects.

In comparison, the advanced-SA participants demonstrated an entirely different pattern. They demonstrated a considerable degree of stability, as the average score of their definite plurals was high (93.28%; SD: 0.14), and the average score of their bare plurals strikingly dropped: (6.72%; SD: 0.14), with a significant difference between both averages:  $t(9) = -7.4703, p = 0.0002$ .

**Figure 1**

*Distribution of Definite and Bare Plurals in Generic and Specific Readings in Sentence Completion Task*

Task



To determine whether the learning context had direct effects on the patterns of protracted instability of the target forms, an independent sample *t*-test was conducted on the two groups' averages of generic definite plurals. It yielded a significant difference:  $t(19) = -3.9943$ ,  $p = .0013$ . Additionally, the advanced-SA participants' average score of generic definite plurals was compared to that of the native speakers group. There was almost no difference between both averages:  $t(19) = -2.1844$ ,  $p = .0452$ . This result is important in the discussion of the role of the learning context in relation to protracted instability of generic definite plurals.

**Table 3***Individual Analysis of Generic Plural Nouns in the Sentence Completion Task*

Performance level	Advanced-AH (N = 10)		Advanced-SA (N = 10)	
	Definite	Bare	Definite	Bare
Very low: 0%-24%	3 of 10	1 of 10	0 of 10	9 of 10
Low: 25%-49%	5 of 10	3 of 10	0 of 10	1 of 10
Mid: 50%-74%	1 of 10	3 of 10	1 of 10	0 of 10
High: 75%-89%	0 of 10	1 of 10	1 of 10	0 of 10
Ceiling: 90%-100%	1 of 10	2 of 10	8 of 10	0 of 10

To scrutinize variation in individual performance on the generic reading condition, averages of definite and bare plurals by the two groups were examined along five performance levels. Results are presented in Table 3. For definite plurals, the majority of the advanced-AH participants (8 out of 10) were in the lowest two ranges. In other words, they were below 49%, not showing much departure from their average score (49.66%). Only one participant performed in the middle range, and one participant performed in the highest range. In comparison, the advanced-SA group demonstrated a different pattern. None of the participants produced definite plurals in the lowest two ranges, and the majority of the participants (9 out of 10) produced definite plurals in the highest two ranges (75%-100%), and one participant was in the middle range. This does not show much departure from the group's average (93.28%).

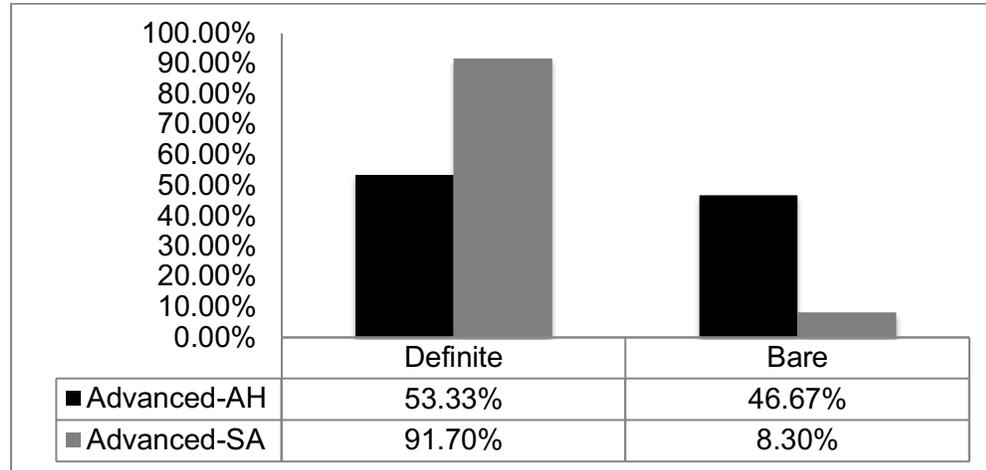
The distribution of bare plurals across the aforementioned performance levels was in contrast to the distribution of definite plurals. In the advanced-AH group, 7 out of the 10 participants produced bare plurals (transferred from their L1 English) in the highest three ranges. The advanced-SA group's minimal protracted instability of the target forms and retreat from L1 effects was supported by the absolute absence of bare plurals in the highest three ranges for the ten participants. Except for only one participant who produced bare plurals in the 25%-49 range, the rest of the participants produced bare plurals in the lowest range. Overall, the individual analysis of target-like (definite plurals) and the non-target-like (bare plurals) forms confirmed the general trends of the group results of the two L2 advanced groups. It further sheds light on the patterns of residual optionality and retreat from L1 effects.

### *Elicited oral narrative*

Overall, the performance of the two advanced groups in the oral narrative task was not different from their performance in the sentence completion task. As illustrated in Figure 2, the advanced-AH participants clearly fluctuated between bare and definite plurals. Specifically, the average score of their definite plurals was slightly beyond chance (53.33%; SD: 0.10), and their average score of bare plurals was slightly lower (46.67%; SD: 0.10), with no difference between both averages in a paired-samples  $t$ -test:  $t(9) = -0.6030$ ,  $p = 0.5631$ .

**Figure 2**

*Distribution of Definite and Bare Plurals in Generic Reading in Oral Narrative Task*



The advanced-AH group’s fluctuation between non-target bare plurals and target definite plurals is illustrated by the two representative excerpts in (11) and (12) by the same participant from this group.

- (11) fii raʔjii ʔaðʕun ʔanna **mudun** kabiira  
in my opinion I think that **cities.bare** big  
ʔaḥsan min **mudun** sʕayiira liʔanna fii  
better than **cities.bare** small because in  
**mudun** kabiira hunaaka ʔal-ʔašxaasʕ ʔakəar  
**cities.bare** big there are people more  
ʔaydʕan fii **mudun** kabiira hunaaka matʕaaʕim ʔakəar  
Also, in **cities.bare** big there are restaurants more

“In my opinion, I think big **cities** are better than small **cities** because in big **cities** there are more people. Also, in big **cities** there are more restaurants.”

(12)	fii	raʔjii,	ʔaðʕun	ʔanna	<b>ʔal-ʔatʕfaal</b>	juḥibbuuna	<b>ʔal-...um</b>
	in	my	I think	that	<b>def-</b>	like	<b>the-...um</b>
		opinion,			<b>children</b>		
	<b>ʔal-ʔalwaan.</b>	la	ʔaʕrif.	li-ʔanna	mumkin	li-ʔanna	<b>ʔal-ʔalwaan</b>
	<b>def-colors.</b>	I do not	know.	because	possibly	because	<b>def-colors</b>
	muəiira	wa	<b>ʔal-ʔatʕfaal</b>	juḥibbuuna	ʔal-ʔaʕyaaʔ	muəiira	
	exciting	and	<b>def.children</b>	like	things	exciting	

“In my opinion, I think **children** love **colors**... [pause] um, I do not know, because...[pause] um. May be because **colors** are exciting and **children** love exciting things.”

The extended context from which excerpt (11) was taken provides a generic statement, as to whether big cities in general are better than small cities, followed by further elaboration. Consistently, the participant produced four tokens of the bare plural for *cities* in a context in which definite plurals should have been used. In contrast, in (12) the same participant produced four tokens of the target definite plurals for *children* and *colors*. The context required the participant to make a generic statement, as to whether children like colors, followed by further details. Two definite tokens of the noun plural for *children* and two definite tokens of the noun plural for *colors* were produced by the same participant.

The instability of the target definite plurals was further confirmed by the fluctuation between bare and definite plurals in the same context by some participants in

the advanced-AH group. The excerpts in (13) and (14) clearly illustrate this, as they were produced by the same participant.

The extended excerpt from which (13) was taken provided a generic statement as to whether children love colors and why. The participant produced two tokens of the target noun plural for *children*. Interestingly, one token was bare, and the other one was definite in the same context. Also, they produced a definite token and a bare token of the plural noun for *colors*. This fluctuation was further confirmed in excerpt (14) in which the participant was asked to provide a generic statement as to whether translators need to study more languages nowadays. Two tokens (bare and definite) were produced in the same context.

In comparison to the advanced-AH group, the advanced-SA participants, consistently with their performance in the sentence completion task, did not show fluctuation between definite and bare plurals. Their average score of definite plurals was 91.70%% (SD: 0.07), and the average score of their bare plurals was strikingly low: 8.30% (SD: 0.07), with an expected difference between both averages:  $t(9) = -15.0000, p = 0.0000$ . The stability of the target definite plurals is illustrated by two representative excerpts in (15) and (16) by the same participant from this group.

(13)	naʕam	ʔaðʕun	<b>ʔal-ʔawlaad...</b>	<b>ʔawlaad</b>	juħibbuuna
	Yes	I think	<b>def-children...</b>	<b>children.bare</b>	like
	<b>ʔalwaan</b>	liʔannahum	jušaahiduunahaa	kul	jawm

<b>colors.bare</b>	because they	watch them	every	day		
<b>ʔal-ʔalwaan</b>	taxtalif	min	šajʔ	lišajʔ	fii	ʔal-bajt
<b>def-colors</b>	differ	from	something	to something	at	home

“Yes, I think **children** love **colors**. They like **colors** because they watch them every day. **Colors** differ from one thing to another.”

(14)

biʔalnisbati	lii,	ʔaðʕun	<b>ʔal-</b>	biʔalnisbati	lii
According	to me,	I think	<b>mutardʒimuun...</b>	according to	me
ʔaðʕun	<b>mutardʒimuun</b>	jadʒib	ʔan	jadrusuu	
I think	<b>translators.bare</b>	have	to	study	
ʔakəar	min	luʔa	liʔanna	fiih	
more	than	one language	because	there are	
ʔakəar	min	luʔa	fii	ʔal-ʕaalam	
more	than	one language	in	the world	

“For me, I think **translators** have to study more than one language because there is more than one language in the world.”

(15)

hunaaka	ʔixtilaafaat	kaəiira	bajna		
There are	differences	big	between		
<b>ʔal-mudun</b>	ʔal-kabiira	wa	<b>ʔal-mudun</b>	ʔal-sʕayiira	
<b>def-cities</b>	def-big	and	<b>def-cities</b>	def-small	
wa	min	ʔal-sʕaʕb	ʔan	ʔaquul	
and	it is	difficult	to	say	
ʔiðaa	kaana	jaʕnii	waaḥid		
whether	was	like	one		
min	ʔal-ʔanwaaʕ	min	ʔal-mudun		
of	the kinds	of	cities		
ʔaḥsan	min	ʔal-ʔanwaaʕ	ʔal-ʔuxraa		
better	than	kinds	other		

“There are big differences between big **cities** and small **cities**, and it’s difficult to say whether one of these kinds of cities is better than the other.”

(16)	ʔað <sup>s</sup> un I think	ʔanna that	<b>ʔal-talaamiið</b> <b>def-students</b>	ʔaw and	<b>ʔal-ʔat<sup>s</sup>faal</b> <b>def-children</b>
	juhibbuuna like	<b>ʔal-ʔalwaan.</b> <b>colors.</b>	ʔað <sup>s</sup> un I think	ʔinnu that	jaʔnii there are
	kaøiir a lot	min of	ʔal-ʔasbaab reasons	tijji coming	faqat <sup>s</sup> only
	min from	ʔal-øaqaafa culture	wa-lajsa and not	faqat <sup>s</sup> only	min from
	ʔal-sajkuluudzijja psychology	ʔinda of	ʔal-t <sup>s</sup> ifl the child		

“I think **students** or **children** like **colors**. I think that a lot of reasons come from culture and not from the child’s psychology.”

For the two advanced learner groups, to determine the difference between the patterns of their in/stability of the target definite plurals and retreat from L1 effects, the difference between their averages of definite plurals was calculated in an independent *t*-test. It returned a significant difference:  $t(19) = 8.2952, p = .0000$ .

***Research question 2: The learning context factor in depth***

Results of the two prompted production tasks answered the first research question. They showed that the advanced-SA participants demonstrated a remarkably more stable pattern in producing generic definite plurals than the advanced-AH participants. To a great extent, the advanced-SA participants opted for target-like definite plurals. To address the second research question, which is whether in/stability patterns of the target-

like plurals produced by the two advanced L2 groups are associated with the input conditions in their respective learning contexts, the contact hours (in-class and out-of-class) in the LCP survey per week for each group were calculated. Results are reported in Table 7.

**Table 7**

*Distribution of In-Class and Out-of-Class Language Contact Averages per Week*

	Advanced-AH		Advanced-SA	
	Mean	SD	Mean	SD
In-class contact				
Global	5.00	0.79	19.43	1.31
○ Speaking	1.43	0.58	4.85	0.89
○ Reading	1.20	1.30	5.80	1.21
○ Writing	1.07	0.31	2.78	0.39
○ Listening	1.30	0.31	6.00	0.78
Out-of-class contact				
Global	13.61	1.23	76.41	4.36
○ Speaking	4.05	0.72	18.85	1.95
○ Reading	3.88	1.05	21.71	3.54
○ Writing	2.33	0.70	12.14	1.57
○ Listening	3.33	1.58	14.71	1.88

For the weekly average of in-class contact hours, the advanced-SA group reported more hours in Arabic across the four skills than the advanced-AH group with a significant difference between both:  $t$ -test:  $t(19) = -14.8447$ ,  $p = 0.0000$ . For reading, for example, the

pedagogical approach used with the advanced-AH group focused on a limited number of texts, given an approach that focused on teaching the four language skills in an integrated fashion for five hours a week. In comparison, the advanced-SA group studied a wide array of reading texts about contemporary Arab culture offering plentiful amounts of input in an independent reading class for 5-6 hours a week. In a similar vein, the advanced-SA group reported a significantly higher out-of-class contact average than that of the SA group across the four language skills:  $t(19) = -17.6280$ ,  $p = 0.0001$ . They varied significantly: 76.41 hours compared to only 13.61 hours.

The advanced-AH group mostly used Arabic out of class to do homework assignments. In contrast, the advanced-SA group reported a myriad of occasions on which they were in active contact and use of Arabic out of class. These involved conversing on a daily basis with native Arabic speakers, such as service personnel in service encounters. Moreover, they reported deliberately using what they were taught in class (e.g., grammar, vocabulary, and expressions) in speaking with native speakers out of classroom. They also reported reading newspaper articles and literary texts on a daily basis (a minimum of two hours) out of class. Finally, they were required to listen to Arabic in the media for at least two hours a day, five days a week.

Another salient feature of the advanced-SA group's learning context was a pedagogical approach that integrated reading and writing. They reported writing in Arabic out of class for a minimum of five days a week. In addition to the short papers that

they wrote in reaction to the reading topics covered in class, they were required to submit a 300-400 word essay on a weekly basis. These essays were “dissected sentence per sentence” and non-native-like stylistic, pragmatic, and grammatical errors were corrected. No single error “was left” behind.

## **Discussion**

The present study examines the effects of learning context on the patterns of persistent instability in a syntax-semantics interface property, namely generic definite plurals in L2 Arabic. For the first research question, the results of the two elicitation tasks show that the advanced-SA group exhibited a significantly more stable pattern than the advanced-AH group in their production of target-like generic definite plurals. Whereas the SA group consistently produced target-like generic definite plurals, the AH group fluctuated between bare and definite plurals. For the second research question, the results of the language contact profile demonstrates that the advanced-SA group reported significantly more in-class and out-of-class contact hours than the advanced-AH group.

The results of the advanced-AH group suggest that the protracted instability of the target forms continues to persist at advanced proficiency levels in syntax-semantics interface properties in cases of impoverished input. This result is consistent with the predictions of the original proposal of the Interface Hypothesis (Tsimplici & Sorace, 2006). The integration of semantic knowledge (generic vs. specific) with syntactic knowledge

(bare vs. definite plurals) remained far from optimal and gave rise to protracted instability. As a consequence, the advanced-AH group fluctuated between bare and definite plurals. In the AH learning context, the input did not provide definite plurals that encode generic readings with sufficient frequency. This lack of quality input is evidenced by how the semantics of definite plurals are presented in the textbook used by the advanced-AH students, which is *Al-Kitaab fii Ta'allum al-'Arabiyya* (Brustad et al. 2001). When plural nouns are presented in the early lessons of Part 1, the focus is mostly on introducing students to the numerous plural patterns and how they vary according to the inherent linguistic features of nouns such as gender (masculine vs. feminine) and humanness (human vs. non-human). For example, on pages 27-32, many plural noun patterns with examples under each are listed. In lesson 3 on pages 50-51, more plural noun patterns are introduced in the bare form, but no reference is made to the generic-specific distinction. The second and the third parts of *Al-Kitaab* do not provide follow up lessons on the semantics of definite and bare plurals. The series does not explain that definite plurals do not always encode specific reference. In the absence of sufficient quality input that specifically shows the interpretive possibilities of bare and definite plurals, non-target-like L1 forms and the target L2 forms continue to co-exist.

Sorace (2005, 2006) claims that protracted instability is “real” in the verbal domain in the same context in interface properties. In the present study, the advanced-AH group continued to exhibit evidence of fluctuation between bare plurals (transferred from L1)

and target definite plurals in the same context. There was delayed alignment of L1 forms with the target language forms. The competing input from frequent use of L1 in the at-home learning context likely led to persistent difficulty in accomplishing this alignment. Due to a dearth of input, the advanced-AH group could not completely “unlearn” their L1 bare plurals and restructure their grammar using the target definite plurals. This suggests that the preemption of the L1 forms in properties pertaining to the syntax-semantics interface is less likely to occur in learning settings that provide impoverished input in advanced proficiency levels. In comparison, results of the advanced-SA group exhibited a different pattern. The competing non-target-like bare plurals were inhibited in their spontaneous oral production to a great extent. It is very likely that input flooding, which is characteristic of language learning in immersion settings abroad, played a crucial role in suppressing the persistent non-target-like L1 forms and further maintained the target-like forms.

Although this input explanation in the SA setting accounts for how the advanced-SA group acquired the semantic subtleties of definite plurals, it does not explain how they unlearned bare plurals as a grammatical option in generic readings. This task is complicated by the fact that bare plurals are grammatical in Arabic in the preverbal position, but they are assigned an existential reading. How they accomplished this unlearning task remains an interesting research question. A possible explanation is that they were able to utilize indirect negative evidence or the absence of bare plural nouns in

generic reading contexts, à la Ionin and Montrul (2010). These authors speculate that advanced Spanish-speaking learners might have had access to indirect negative evidence in the acquisition of English bare plurals. Following this line of explanation, it is possible that the advanced-SA learners in the present study utilized the non-occurrence of generic bare plurals in the preverbal position as a cue in the input to map a generic reading only to definite plurals. This indirect negative evidence requires inductive inference and making generalizations, which entail taking specific positive evidence into account. The advanced-AH learners were not able to accomplish this unlearning task because this specific type of evidence was not established to test the hypothesis that bare plurals were not a possible option in Arabic generic readings. Only advanced learners who were taught in an immersion setting in the target culture for an extended period of time established this evidence.

The results regarding the marginally protracted instability of definite plurals by the advanced-SA group in this study raise another question about the internal cognitive resources that may have enabled them to be attentive to generic definite plurals (whether through indirect negative evidence or direct positive evidence). A plausible explanation comes from the Threshold Hypothesis (Lafford, 2006). This hypothesis explains the gains in grammatical subtleties in the SA setting in terms of internal cognitive resources such as working memory and attention span. This explanation is gaining ground due to results of research that approaches gains in SA contexts from a psycholinguistic perspective (e.g.,

Harrington & Sawyer, 1992; Hdstijn & Bossers, 1992; Sunderman & Kroll, 2009; Tokowicz, Michael, & Kroll, 2004). The picture that emerges from this research is that advanced-SA learners are better prepared to make greater gains in grammatical subtleties because their solid lexical and grammatical base leaves more cognitive space to attend to the subtle grammatical features that do not have much communicative value or disrupt comprehension, but are often functional in nature. In the present study, it is likely that the advanced-SA group did not exhaust too many attentional resources on the fundamental lexical and grammatical properties of Arabic that they already established before their SA experience. Having automatized the basic lexical and grammatical structures, they were better able to attend to the intricacies of L2 constructions that pertain to the interfaces, such as article semantics. Although the advanced-AH group in this study developed a solid lexical and grammatical base that may have left cognitive space to the subtleties of L2 grammar in their formal classroom setting, the impoverished L2 input and the continuous use of their L1 forms may have resulted in having two competing options in their interlanguage systems.

A central issue in discussions regarding the IH is whether the learnability problems that arise in interface properties are internal or external with regard to learners' representations of these properties. The external account considers the processing difficulties that arise while integrating knowledge from internal or external domains (Sorace, 2005). In the present study, the two oral spontaneous instruments vary in their

processing cost. Oral narrative is taken to be more costly than sentence completion. The fact that the overall pattern of performance was not significantly different in the two tasks may be taken as evidence of how genericity is encoded in plural nouns. The instability of the generic definite plurals by the advanced-AH group can be seen as a manifestation of their indeterminate representations. Along a similar vein, the stability in the advanced-SA group's performance was a manifestation of their determinate "native-like" representations. These results suggest that attaining near-native mental representations is possible in learning settings where the active use of the L2 is maintained for an extended period of time. They also support the position that a native-like steady state is possible in syntax-semantics interface properties (Slabakova & Montrul, 2003; Valenzuela, 2006), and it is constrained by input conditions (e.g., usage and exposure) in immersion settings (Sorace, 2005, 2011).

### **Conclusion, limitations, and future research**

To conclude, the results of the AH group in this study confirm the prediction of the original proposal of the Interface Hypothesis that grammar properties requiring the integration of syntactic knowledge with knowledge from other domains such as lexical semantics continue to present protracted instability at advanced proficiency levels. Syntax-semantics interface properties in which L1 and L2 forms are in conflict continue to be vulnerable in advanced proficiency levels. Results of the advanced-SA group suggest

that immersion programs combining structured formal instruction and continuous exposure and active use of the target language contribute to minimizing protracted instability in syntax-semantics interface properties. The learning context is suggested to play a crucial role in stabilizing target-like L2 forms and preempting competing non-target-like L1 forms in interface properties. Retreat from L1 effects at the interfaces is possible at advanced proficiency levels in immersion settings. As Sorace (2011) puts it, better understanding of the underpinnings of the interface phenomena requires “interdisciplinary work that benefits from a wider range of methodologies and perspectives” (p.27). Putting the predictions of the IH in learning settings with different characteristics is a step forward to examine the interaction between linguistic and non-linguistic factors in the acquisition of interface properties.

A number of methodological and conceptual questions arise from this study. The limited number of advanced participants in each group calls for a larger scale study on protracted instability of generic definite plurals as an interface property in different learning contexts. Although the patterns of in/stability of generic definite plural nouns may be taken as manifestations of the in/determinate mental representations of the target construction, a further study using an online processing methodological design is needed to provide neuropsychological evidence that supports this account. Furthermore, the speculation that the advanced-SA group may have been attentive to the definite generic plurals in the input due to their better internal cognitive resources needs to be scrutinized

in a separate study. Probing into working memory and measuring reaction times will provide evidence that may further support this account. Future work will need to examine the systematic changes in the patterns of protracted instability in article use semantics, as well as other interface phenomena, in search for a threshold beyond which this optionality vanishes. Patterns of optionality in interface properties and the gradual increase of the target forms need to be examined in a longitudinal study in advanced adult L2 acquisition. A last research study will need to examine the association between advanced learners' protracted instability patterns and their metalinguistic knowledge. Utilizing qualitative instruments such as think-aloud protocols (Ericsson & Simon, 1993; Roehr, 2006, 2007) is expected to probe into advanced learners' mental representations of how genericity is encoded in noun phrases.

## References

- Alhawary, M. T. (2009). *Arabic second language acquisition of morphosyntax*. Yale University Press.
- Al-Ani, S. (1972–1973). Features of interference in the teaching of Arabic composition. *An-Nashra*, 5, 3–13.
- Azaz, M. (2014). *Acquisition of form-meaning mapping in L2 Arabic and English noun phrases: A bidirectional framework* [Ph.D. dissertation, University of Arizona].
- Azaz, M. (2016). Crosslinguistic effects in L2 acquisition: The case of Arabic determiner phrase. *Al-'Arabiyya: Journal of the American Association of Teachers of Arabic*, 49(1), 1-24.
- Azaz, M., & Frank, J. (2017). Bidirectional cross-linguistic influence in late bilingualism: Evidence from the container-content relation. *Linguistic Approaches to Bilingualism*. doi 10.1075/lab.15012.aza.
- Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language & Linguistic Theory*, 25(4), 657-689.
- Belletti, A., & Leonini, C. (2004). Subject inversion in L2 Italian. *EUROSLA Yearbook*, 4, 95-118.
- Borgonovo, C., De Garavito, J. B., & Prévost, P. (2008). Methodological issues in the L2 acquisition of a syntax/semantics phenomenon: How to assess L2 knowledge of

- mood in Spanish relative clauses. In J. Bruhn de Garavito & E. Valenzuela (Eds.), *Selected proceedings of the 10<sup>th</sup> Hispanic Linguistics Symposium* (pp. 13–24). Cascadilla.
- Brustad, K., Al-Batal, M., & Tunisi, A. (2001). *Al-Kitaab fii ta'allum Al-'Arabiyya: A textbook for Arabic*. Georgetown University Press.
- Carroll, J. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals*, 1, 131-151.
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6, 339–405.
- Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26, 229–250.
- Collentine, J., & Freed, B. (2004). Learning context and its effects on second language acquisition: Introduction. *Studies in Second Language Acquisition*, 26, 153–171.
- Cuza, A., & Frank, J. (2011). Transfer effects at the syntax–semantics interface: The case of double que questions in heritage Spanish. *The Heritage Language Journal*, 8, 66–89.
- Cuza, A., Guijarro-Fuentes, P., Pires, A., & Rothman, J. (2012). The syntax-semantics of bare and definite plural subjects in the L2 Spanish of English natives. *International Journal of Bilingualism*, 17, 632–52.
- Dayal, V. (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy*, 27, 393–450.

- de Garavito, J. (2013). Semantic features in L2 Spanish: A response to Cuza, Guijarro-Fuentes, Pires, and Rothman. *International Journal of Bilingualism*, 17(5), 653-656.
- DeKeyser, R. M. (1991). Foreign language development during a semester abroad. In B. Freed (Ed.), *Foreign language acquisition and the classroom* (pp. 104-118). Heath and Company.
- Dekydtspotter, L., & Sprouse, R. (2001). Mental design and (second) language epistemology: Adjectival restrictions of wh-quantifiers and tense in English-French interlanguage. *Second Language Research*, 17, 1-35.
- Dekydtspotter, L., Sprouse, R. A., & Anderson, B. (1997). The interpretive interface in L2 acquisition: The process-result distinction in English-French interlanguage grammars. *Language Acquisition*, 6(4), 297-332.
- Dekydtspotter, L., Sprouse, R., & Swanson, K. (2001). Reflexes of mental architecture in second-language acquisition: The interpretation of combien extractions in English-French interlanguage. *Language Acquisition*, 9, 175-227.
- Ericsson, A., & Simon, H. (1993). *Protocol analysis: Verbal reports as data*. MIT Press.
- Fassi Fehri, A. (2004). Nominal classes, reference, and functional parameters, with particular reference to Arabic. *Linguistic Variation Yearbook*, 4, 41-107.
- Freed, B. (1995). Second language acquisition in a study abroad context. John Benjamins.
- Freed, B. F., Dewey, D. P., Segalowitz, N. S., & Halter, R. H. (2004). The language contact profile. *Studies in Second Language Acquisition*, 26(02), 349-356.

- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26(02), 275-301.
- Flege, J. E., Mackay, R. A, & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, 23(4), 567-598.
- Gabriele, A. (2009). Transfer and transition in the SLA of aspect. *Studies in Second Language Acquisition*, 31(3), 371- 402.
- Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. Erlbaum.
- Greenberg, Y. (2007). Exceptions to generics: Where vagueness, context dependence, and modality interact. *Journal of Semantics*, 24(2),131–167.
- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14(01), 25-38.
- Haznedar, B. (2010). Transfer at the syntax-pragmatics interface: Pronominal subjects in bilingual Turkish. *Second Language Research*, 26(3), 355-378.
- Hertel, T.J. (2003). Lexical and discourse factors in the second language acquisition of Spanish word order. *Second Language Research*, 19, 273–304.
- Hopp, H. (2009). The syntax–discourse interface in near-native L2 acquisition: Off-line and online performance. *Bilingualism: Language and Cognition*, 12, 463–483.

- Huebner, T. (1995). The effects of overseas language programs: Report on a case study of an intensive Japanese course. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 171-193). John Benjamins Publishing Company.
- Hdstijn, J. H., & Bossers, B. (1992). Individual differences in L2 proficiency as a function of L1 proficiency. *European Journal of Cognitive Psychology, 4*(4), 341-353.
- Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition, 12*(1), 3-69.
- Ionin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning, 60*(4), 877-925.
- Ionin, T., Montrul, S., & Crivos, M. (2013). A bidirectional study on the acquisition of plural noun phrase interpretation in English and Spanish. *Applied Psycholinguistics, 34*(03), 483-518.
- Ionin, T., Montrul, S., Kim, J. H., & Philippov, V. (2011). Genericity distinctions and the interpretation of determiners in second language acquisition. *Language Acquisition, 18*(4), 242-280.
- Isabelli, C. A., & Nishida, C., (2005). Development of the Spanish subjunctive in a nine-month study-abroad setting. In D. Eddington (Ed.) *Selected proceedings of the 6<sup>th</sup> Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 78-91). Cascadilla.

- Ivanov, I. (2009). Topicality and clitic doubling in L2 Bulgarian: A test case for the interface hypothesis. In M. Bowles, T. Ionin, S. Montrul, & A. Tremblay (Eds.), *Proceedings of the 10<sup>th</sup> Generative Approaches to Second Language Acquisition Conference*, (pp. 17–24), Cascadilla.
- Krifka, M., Pelletier, F., Carlson, G., ter Meulen, A., Chierchia, G., & Link, G., (1995). Genericity: An introduction. In G. Carlson & F. Pelletier (Eds.), *The generic book*. (pp.1–124), University of Chicago Press.
- Lafford, B. A. (2006). The effects of study abroad vs. classroom contexts on Spanish SLA: Old assumptions, new insights, and future research directions. In C. A. Klee & T. L. Face (Eds.), *Selected proceedings of the 7<sup>th</sup> conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 1-25). Cascadilla.
- Lapkin, S., Hart, D., & Swain, M. (1995). A Canadian interprovincial exchange: Evaluating the linguistic impact of a three-month stay in Quebec. In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 67- 94). John Benjamins.
- Llanes, À., Mora, J. C., & Serrano, R. (2016). Differential effects of SA and intensive AH courses on teenagers' L2 pronunciation. *International Journal of Applied Linguistics*, 27(2), 470-490.
- Longobardi, G. (2001). How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Natural Language Semantics*, 9, 335-69.

- Lozano, C. (2006). Focus and split-intransitivity: The acquisition of word order alternations in non-native Spanish. *Second Language Research*, 22, 145-187.
- Marqués-Pascual, L. (2011). Study abroad, previous language experience, and Spanish L2 development. *Foreign Language Annals*, 44(3), 565–582.
- Masuda, K. (2011). Acquiring interactional competence in a study abroad context: Japanese language learners' use of interactional particle Ne. *Modern Language Journal*, 95(4), 519-540.
- Montrul, S., & Ionin., T. (2010). Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition*, 13, 449-73.
- Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite/imperfect contrast in Spanish. *Studies in Second Language Acquisition*, 25, 351-98.
- Rammuny, R. (1976). Statistical study of errors made by American students in written Arabic. *Al-'Arabiyya: Journal of the American Association for Teachers of Arabic*, 9,75-94.
- Rebuschat, P., & Mackey, A. (2013). Prompted production. In C. Chappelle (Ed.), *The encyclopedia of Applied Linguistics*, Vol. 5, (pp. 75-84). Wiley–Blackwell.
- Roehr, K. (2006). Metalinguistic knowledge in L2 task performance: A verbal protocol analysis. *Language Awareness*, 15, 180-198.
- Roehr, K. (2007). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics*, 29, 173-190.

- Rothman, J. (2009). Pragmatic deficits with syntactic consequences? L2 pronominal subjects and the syntax-pragmatics interface. *Journal of Pragmatics*, 41, 951-973.
- Segalowitz, N., & Freed, B. F. (2004). Con-text, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173-199.
- Segalowitz, N., Freed, B., Collentine, J., Lafford, B., Lazar, N., & Díaz-Campos, M. (2004). A comparison of Spanish second language acquisition in two different learning contexts: Study abroad and the domestic classroom. *Frontiers: The interdisciplinary journal of study abroad*, 10, 1-18.
- Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2009). Bilingual children's sensitivity to specificity and genericity: Evidence from metalinguistic awareness. *Bilingualism: Language and Cognition*, 12(02), 239-257.
- Slabakova, R., & Montrul, S. (2003). Genericity and aspect in L2 acquisition. *Language Acquisition*, 11(3), 165-196.
- Slabakova, R., & Ivanov, I. (2011). A more careful look at the syntax-discourse interface. *Lingua*, 121, 637-651.
- Smith, N. V. (1975). On generics. *Transactions of the Philological Society*, 27-48.
- Sorace, A. (2000). Syntactic optionality in non-native grammars. *Second Language Research*, 16, 93-102.

- Sorace, A. (2005). Selective optionality in language development. In L. Cornips, & K. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and the social* (pp. 55-80). John Benjamins.
- Sorace, A. (2006). Gradedness and optionality in mature and developing grammars. In G. Fanselow, C. Féry, R. Vogel, & M. Schlesewsky (Eds.), *Gradience in grammar* (pp. 106-123.) Oxford University Press.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1-33.
- Sorace, A. (2014). Input, timing, and outcomes in a wider model of bilingualism. *Linguistic Approaches to Bilingualism*, 4, 377-380.
- Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22, 339–368.
- Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, 13, 195-210.
- Sunderman, G., & Kroll, J. F. (2009). When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics*, 30(01), 79-99.
- Swender, E., Conrad, D., & Vicars, R. (2012). *ACTFL proficiency guidelines*. American Council for the Teaching of Foreign Languages.

- Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, 7(03), 255-272.
- Tsimpli, I. M., Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30<sup>th</sup> Annual BUCLD* (pp. 653–664). Cascadilla.
- Valenzuela, E. (2005). *L2 ultimate attainment and the syntax-discourse interface: The acquisition of topic constructions in non-native Spanish and English* [Ph.D. dissertation, McGill University].
- Valenzuela, E. (2006). L2 end state grammars and incomplete acquisition of the Spanish CLLD constructions. In R. Slabakova, S. Montrul, & P. Prévost (Eds.), *Inquiries in linguistic development: In honor of Lydia White* (pp. 283-304). John Benjamins.
- White, L. (2011). Second language acquisition at the interfaces. *Lingua*, 121(4), 577-590.