# **Between syntax and discourse:** Topic and case marking in heritage speakers and L2 learners of Japanese and Korean


## Oksana Laleko & Maria Polinsky

**State University of New York at New Paltz | University of Maryland College Park**



▼ **Abstract**

This article examines the knowledge of topic and subject particles in heritage speakers and L2 learners of Japanese and Korean. We assume that topic marking is mediated at the syntax-information structure interface, while subject marking pertains to narrow syntax. In comparing phenomena mediated at different levels of linguistic organization, we provide evidence for the hypothesis that information structure-level phenomena present greater challenges for bilingual speakers than those mediated within syntax. While these results may be interpreted as evidence of generalized interface-related deficits, we show that such a global explanation is not supported. Instead, a more nuanced account is developed, based on the recognition of different types of topic (anaphoric, generic, and contrastive) and different types of subject (descriptive and exhaustive). Under the proposed account, the non-native speakers' deficits follow from three unrelated effects: the status of topic as an interface category, structural complexity, and the memory demands necessary for its interpretation in context.

▼ **Publication history**

Date received: 27 April 2014

Date accepted: 08 April 2015

Published online: 25 March 2016

▶ **Table of contents**

doi: **10.1075/lab.14018.lal** | Published online: 25 March 2016

*Linguistic Approaches to Bilingualism* | ISSN 1879-9264 | E-ISSN 1879-9272



# 1. Introduction

## 1.1 Interfaces: The integration problem

Heritage language and second language (L2) acquisition represent two distinct pathways to adult bilingualism. They differ along a number of dimensions, including the specific circumstances of language exposure and the nature of competence deficits across different language modules. As early bilinguals, heritage speakers begin the acquisition of a socially non-dominant language in a family setting, and in this respect, their experience with the target language compares to that of early L1 learners, at least at the outset of the language acquisition process. The main difference is that heritage speakers' L1 language acquisition is subsequently interrupted by exposure to another language. Adult L2 learners, on the other hand, are late bilinguals whose exposure to the non-dominant language typically takes place in an instructed setting.

Despite these differences in acquisition scenarios, both groups of speakers similarly fall short of exhibiting

complete mastery of the target language. Even at the highest levels of proficiency, both populations remain in a state of near-native attainment, deviating from target native grammars represented by the monolingual baseline. In particular, early and late bilinguals (heritage speakers and L2 learners alike) have been reported to display deficits in the domain of inflectional morphology and narrow syntax, and both groups also seem to have difficulties with discourse-level phenomena. The former problem is manifested in errors or non-target-like performance with case, gender, agreement, verbal aspect, and long-distance dependencies (Benmamoun, Montrul, & Polinsky, 2010; Benmamoun, Montrul, & Polinsky, 2013a, b; Hawkins & Hattori, 2006; Montrul, 2002, 2005; Montrul, Foote, & Perpiñán, 2008; Polinsky, 1997, 2006, 2008a, b, 2011; White, 2003), and the latter problem involves infelicitous linguistic choices in contexts that require discourse tracking or external pragmatic knowledge to resolve apparent contextual optionality (Laleko, 2010; Laleko & Polinsky, 2013; Montrul, 2004; Serratrice, Sorace, & Paoli, 2004; Sorace, 2011; Sorace, Serratrice, Filiaci, & Baldo, 2009).

Research on advanced second language learners and heritage speakers of comparable proficiency has additionally demonstrated that certain discourse properties remain problematic even after (morpho-)syntactic properties are in place. Results obtained from advanced speakers constitute evidence for the "syntax-before-discourse" hypothesis in second language acquisition (Pérez-Leroux & Glass, 1997; Polio, 1995; Rothman, 2007, 2009) and adult L1 attrition (Sorace, 2004, 2011). According to this hypothesis, syntactic competence is acquired sooner and is more immune to attrition than discourse-pragmatic knowledge. While research involving high-proficiency heritage speakers is still relatively scarce, there is some evidence indicating that these speakers have non-target-like mastery of some discourse-related phenomena in the absence of morphosyntactic errors. For example, a study of heritage speakers' knowledge of the Russian aspectual system by Laleko (2010) found that high-proficiency heritage speakers differed from monolingual controls in the use and interpretation of contextually dependent functions of the imperfective aspect, including those related to the marking of old and new information

in discourse, but not on its core sentence-level functions, such as indicating habitual and progressive aspect. On a theoretical plane, a number of researchers have advanced the idea that the discourse component of the grammar may be more "costly" in terms of linguistic encoding and processing than the core of the syntactic computational component (Langacker, 2000; Reuland, 2011). The theoretical model by Reuland (2011) makes a principled distinction between dependencies established in narrow syntax, logical-form (LF) dependencies, and discourse dependencies. Assuming that grammar and parser are tightly linked (the so-called single-system approach; see Phillips, 2013; Lewis & Phillips, 2015), these three types of dependencies are handled by autonomous comprehension mechanisms. Indeed, Reuland (2011) proposes a hierarchy in which the greatest processing costs are associated with discourse dependencies, while syntactic dependencies are hypothesized to be relatively more economical in terms of processing resources. Preference for structural relationships, such as bound-variable dependencies, over discourse relationships (particularly co-referential dependencies) has been well documented in offline and online studies for various populations of speakers, including agrammatic aphasics as well as unimpaired children and adults (Koornneef, 2008; Koornneef, Avrutin, Wijnen, & Reuland, 2011, and references therein).

The specific factors contributing to the apparent asymmetry between (morpho-)syntactic and discourse components of language are still poorly understood. Perhaps one of the most frequently evoked accounts of the "syntax-before-discourse" problem is a proposal referred to in the literature as the Interface Hypothesis, which claims that advanced bilinguals do not display complete linguistic knowledge in specific interface-conditioned domains, i.e., domains that involve the integration of two or more linguistic modules or mapping between linguistic and non-linguistic information. Under the generative view of language as a system of several semi-autonomous modules interacting with each other (Jackendoff, 2002), the process of language acquisition involves learning the individual modules of language along with the principles of interaction among the language modules. These

mappings among linguistic sub-modules, or between language and external cognitive systems, have been claimed to be subject to vulnerability and incomplete attainment (Argyri & Sorace, 2007; Avrutin, 1999; Bos, Hollebrandse, & Sleeman, 2004; Hulk & Müller, 2000; Müller & Hulk, 2001; Sorace, 2004, 2005, 2011; Sorace & Serratrice, 2009; Tsimpli, Sorace, Heycock, & Filiaci, 2004; Tsimpli & Sorace, 2006). A further distinction can be made between external and internal interfaces. This distinction has been supported by findings that reported differential results for the knowledge of phenomena at the interface between syntax and semantics, e.g., generic determiners and focus marking, and phenomena governed by external pragmatic conditions, e.g., topicalization and pronominal anaphora (Sorace, 2011 and references therein). External interfaces, which involve integration of grammatical material with pragmatic and contextual information, appear to be more problematic than internal interfaces, which involve information within the sub-modules of language (Sorace, 2009; Tsimpli & Sorace, 2006).

The Interface Hypothesis has generated considerable attention in the bilingualism literature in the recent years, and a significant number of empirical studies have been carried out to test its predictions in a variety of languages and for various bilingual populations. Work by Sorace and colleagues on the distribution of overt pronominal subjects in null subject languages has provided an important body of data in support of the main tenet of the Interface Hypothesis: domains linked to external contextual information (e.g., the syntax-pragmatics interface) are more unstable than domains that only involve features internal to linguistic representations (e.g., the syntax-semantics interface). Subsequent research has documented similar asymmetries with respect to other linguistic phenomena (Belletti & Leonini, 2004; Lozano, 2006; Valenzuela, 2006). However, other empirical studies have either failed to detect any apparent signs of instability in domains predicted by the Interface Hypothesis to be vulnerable or suggested that such interface-related difficulties are not pervasive (Donaldson, 2011, 2012; Ivanov, 2012; Iverson, Kempchinsky, & Rothman, 2008; Leal Méndez, Rothman, & Slabakova, 2015; Slabakova & Ivanov, 2011; Slabakova, Kempchinsky, & Rothman, 2012). Researchers have also challenged some aspects of the

hypothesis on theoretical grounds, including the imprecise formulation of the proposal, difficulty of distinguishing interface-related phenomena from non-interface-related phenomena (and, by the same token, external interfaces from internal interfaces), difficulty of positioning the hypothesis within specific theoretical models of language architecture, and potentially an overly "restrictive" focus of the proposal, originally formulated only for some bilingual populations and not others (Domínguez, 2013; Montrul, 2011; Rothman, 2009; Rothman & Slabakova, 2011; White, 2011).

Why would interface-related properties in particular be associated with a greater processing difficulty? The explanation embedded in the Interface Hypothesis is that interface-conditioned domains are intrinsically more difficult in terms of processing due to their integrative nature, i.e., the need to integrate syntactic knowledge with information from other cognitive domains. We will refer to this as the *integration problem*. Under this account, constructions at the interfaces are predicted to be problematic by virtue of the additional processing demands necessary for establishing connections across domains.

## 1.2 Interfaces: Introducing new considerations

Because of the inherent strength and appeal of the Interface Hypothesis, researchers have focused mainly on integrative problems of the interfaces. However, processing difficulties observed for interface-related phenomena may also stem from factors not linked directly to the integration problem as described above. The integration problem could co-exist with other reasons underlying processing difficulties, resulting in a certain "division of labor" according to which specific difficulties faced by non-native speakers are due to several independent factors. Understanding the nature of these difficulties requires a more nuanced and multifaceted approach. In what follows, we outline some components of such an approach by considering two specific dimensions of the "syntax-discourse" asymmetry.

### 1.2.1 *Structural complexity*

The first dimension has to do with the greater *structural complexity* of elements located at the highest levels of sentential structure. Under this view, it is not the need to integrate information from multiple linguistic and non-linguistic domains that serves as a potential source of greater processing difficulties associated with a given construction, but the syntactic structure of the constituent that houses the construction, namely its size and degree of embedding. The interaction between syntactic complexity and processing complexity is a long-standing issue in the literature on the language processing system (Ferreira & Clifton, 1986; Frazier, 1985; Stevenson & Merlo, 1997), and the idea that structural information affects sentence processing is no longer controversial despite the fact that consensus has not been reached on the exact mechanism of this interaction. Assuming a hierarchical organization in (1) below, the approach that we explore here maintains that constructions located in higher clausal projections are more structurally complex and will be associated with greater difficulty than constructions housed in lower syntactic projections, regardless of their interface status.

Basic assumptions about sentence structure (Chomsky, 1995; Rizzi, 1997) yield a hierarchical representation consisting of three universal phrases: the verb phrase (vP), the finite inflectional domain (TP), and the clausal periphery (CP):

(1) $[_{CP}\ldots\ [_{TP}\ldots\ [_{vP}\ldots\ ]]]$

Syntactic information is mediated at the level of the vP and TP projections. We assume that this is the domain that houses operations contained within narrow syntax, i.e., the inflectional level at which morphological relations, such as case licensing, are established. The topmost projection, the CP, is where information encoded at vP and TP levels is linked to discourse (Rizzi, 1997, and much subsequent work). Discourse-related categories such as topic

and focus as well as the illocutionary force of the utterance are represented as unique projections within the split CP (after Rizzi, 1997):

(2) $[_{ForceP} \ldots [_{TopP} \ldots [_{FocP} \ldots [_{TopP} \ldots [_{FinP} \ldots [_{TP} \ldots]]]]]]$

As a natural outcome of such hierarchical organization, constructions that employ projections within the split CP, such as a TopP or a FocP, require building and interpreting more structure than constructions housed within the lower projections, such as the TP. Greater structural complexity predicts greater computational effort associated with larger segments, additional embeddings, and longer dependencies, potentially resulting in increased difficulty in interpretation and production. Previous research has identified several populations of language speakers, including early L1 learners, children with Specific Language Impairment, Broca's aphasia patients, and L2 learners, who have been shown to exhibit systematic difficulties with features located at the CP level while remaining target-like with respect to constructions on the lower syntactic levels (Avrutin, 1999; Platzack, 2001). Most of the empirical support for the Interface Hypothesis encompasses the types of phenomena that involve, in syntactic terms, the highest level of sentential structure (the C-domain). We must consider whether it is possible that difficulty with these phenomena relates to their position within the clausal architecture, rather than to their inherent interface status *per se*.

This approach differs from some previous syntactic accounts of the so-called interface-related difficulties, in which complexity has been discussed primarily in terms of the integrative nature of these phenomena, rather than their syntactic structure. For example, Sorace (2005) appeals to the notion of syntactic complexity to describe properties of certain constructions and syntactic operations in the following way: constructions that require the integration of syntactic knowledge with information from other domains are more complex than constructions that only require syntactic knowledge, and syntactic operations that are obligatorily required are less complex than

those occurring optionally as a result of semantic or pragmatic choices (p. 69). Notice that this account centers around two factors, both ultimately related to the need to incorporate information from various domains, i.e., the integration problem. In contrast, the syntactic account we would like to pursue here investigates structural complexity relative to the syntactic position of a constituent as a possible factor contributing to the asymmetry between phenomena that belong to the realm of narrow syntax and phenomena governed by both syntactic and discourse-level factors.

### **1.2.2** *Memory load*

Due to the costs of storing more information in memory during syntactic computation, the difficulty of discourse-level elements may also stem from the greater *memory demands* required for successful interpretation of elements characterized by context-resolvable optionality. Greater computational demands have been discussed among the primary sources of processing delays associated with interface vulnerability (Sorace, 2011), suggesting a direct link between processing difficulties associated with certain linguistic phenomena and their interface status. However, it is important to recognize that processing difficulties arising from an excessive load on working memory are not *a priori* limited to the interface domains and have been documented for constructions mediated entirely within narrow syntax, i.e., those that do not necessarily involve integration of information from different domains (O'Grady, 2011). Memory constraints have been argued to affect sentence processing in domain-internal areas of language structure, including long-distance and nested dependencies mediated entirely within syntax (Chomsky & Miller, 1963). Thus, the integration problem inherent in computing interface-based structures can be distinguished on principled grounds from the broader issue of processing efficiency, which can vary for different types of phenomena both in interface-conditioned as well as non-interface-conditioned domains.

For our purposes, we will assume that some entities may be associated with greater memory demands if their

construal and interpretation requires *referential linking* to some other material. It is not the inherent interface status of a particular phenomenon but the amount of cognitive effort and memory demands involved in its computation which will be taken to be a potential source of increased difficulty and less target-like performance in some bilingual populations.

Following the discussion of the relevant linguistic contrasts in the next section, we outline some ways in which our data may bear on the two accounts discussed above. The goal of this paper is not to argue for one of these two possibilities, a task that would require a battery of experimental tests ideally involving a set of distinct linguistic phenomena, but rather to explore the question of how structural complexity and memory demands can be teased apart as distinct factors contributing to the observed asymmetries between grammatical and discourse-related knowledge in bilingual populations.

Having identified the factors that can contribute to differential processing difficulty, we are now ready to compare the predictions of the three possible accounts of the 'syntax-before-discourse' problem: an interface-based explanation, which attributes greater problems with discourse-level elements to their inherent interface status; a syntactic account, which examines clausal architecture as a source of computational delays; and an account that takes as its main point of departure the demands on short-term memory necessary for reference tracking and resolving contextual optionality. The latter is sometimes bundled together with interface issues under the rubric of the "integration problem" (cf. Sorace, 2011); we are particularly interested in separating the two because that would allow us to differentiate between the integration needs imposed by linguistic structure *per se* and the more general cognitive demands on memory and identification.

The empirical data for this comparison come from two topic-prominent languages (Li & Thompson, 1976), Japanese and Korean. The consideration of these languages will allow us to compare grammatical case marking, which is mediated within narrow syntax, and the marking of information structure, construed within the discourse

component of the grammar.

## 2. Topic and subject marking in Japanese and Korean

### 2.1 The distribution of topic and subject in Japanese and Korean: Basic generalizations

Japanese and Korean are well known for their explicit morphological expression of subjects and topics. Grammatical subject is encoded by the nominative case particles *-ga* (Japanese) and *-i/-ka* (Korean), whereas the particles *-wa* (Japanese) and *-nun/-un* (Korean) mark topics.

Consider the following examples:[1]

(3) a. Sakana-wa tai-ga oisii. *Japanese*
fish-TOP red snapper-NOM delicious
'Speaking of fish, red snapper is delicious.'

b. Sayngsen-un yene-ka massissta. *Korean*
fish-TOP salmon-NOM delicious
'Speaking of fish, salmon is delicious.'

Research on topics in Japanese and Korean traditionally distinguishes between *thematic* and *contrastive* topics (Choi, 1999; Kuno, 1972, 1973; Kuroda, 1965, 2005; Shibatani, 1977). Thematic topics always appear in a clause-initial position. Based on their interpretive properties, they are further subcategorized into two types: *anaphoric* topics, linked to prior discourse via a linguistic or contextual antecedent, as in (4b), and *generic* topics, which refer to a general class of entities, as in (5).

(4) a.

Watasi-wa sengetsu hazimete tyuugoku-o otozure-mas-ita. *Japanese*
1sg-TOP last month first time China-ACC visit-POLITE-PST.DECL
'I visited China for the first time last month.'
Ima tyuugoku-wa itiban sukina kuni-desu.
now China-TOP first favorite country-be.PRS.DECL
'Now, China is my favorite country.'

b. Na-nun cinan tal-cheumulo cwungkwuk-ul pangmwunhay-ss-ta. *Korean*
1sg-TOP last month first time China-ACC visit-PST-DECL
'I visited China for the first time last month.'
Cikum cwungkwuk-un kacang cohaha-nun nala-ita.
now China-TOP first like-ADN country-be.DECL
'Now, China is my favorite country.'

(5) a. Kami-wa ki-kara tuku-rare-masu. *Japanese*
paper-TOP tree-from make-PASS-PRS.DECL
'Paper is made from trees.'

b. Congi-nun namwu-eyse mantule ci-n-ta. *Korean*
paper-TOP tree-from make PASS-PRS-DECL
'Paper is made from trees.'

Unlike thematic topics, whose main functions relate to reference and discourse tracking (either through theme maintenance associated with anaphoric topics or through theme creation associated with generic topics), contrastive topics have a distinct function of signaling a special contrastive relationship between two or more elements within a set of alternatives, as illustrated in (6) below. The contrastive relationship can be explicit, when

both members of the set are present in discourse (as shown in (6)), or implied, when only one member of the contrastive relationship is mentioned (as in (7) below).

(6) a. Watasi-wa hudan hoogaku-wa kikimasu ga yoogaku-wa kiki-masen. *Japanese*
1sg-TOP usually Japanese music-TOP listen but Western music-TOP listen-NEG
'I usually listen to Japanese music, but I do not listen to Western music.'

b. Na-nun pothong hankuk umak-un tut-ciman seyang umak-un tut-ci ahn-nun-ta. *Korean*
1sg-TOP usual Korean music-TOP listen-although Western music-TOP listen-INF NEG-PRS-DECL
'I usually listen to Korean music, but I do not listen to Western music.'

The two different functions of the topic marker, thematic and contrastive, are identical in terms of form in both languages; however, they are distinguished by means of prosody and syntactic behavior (Kuno, 1973; Maynard, 1980; Nakanishi, 2001). Prosodically, contrastive topics carry emphatic stress, while thematic topics are associated with neutral intonation. The syntactic distribution of contrastive topics is less constrained in that contrastive topics can occur in both root and subordinate clauses, whereas thematic topics are largely restricted to root clauses. As a result, topic-marked entities occurring in embedded clauses are interpreted contrastively rather than thematically:

(7) a. Taroo-wa [Hikari-wa kirei-da-to] omou. *Japanese*
Taroo-TOP Hikari-TOP beautiful-be-COMP think
'Taroo believes that Hikari [as opposed to someone else] is beautiful'

b. Chelswu-nun [Mina-nun alumtap-ta-ko] sayngkakhan-ta. *Korean*
Chelswu-TOP Mina-TOP beautiful-be-EMBEDDING think-DECL
'Taroo believes that Hikari [as opposed to someone else] is beautiful'

Japanese and Korean both exhibit flexible word order for verbal arguments, and therefore cannot rely solely upon word order to assign grammatical relations as English does. Instead, both languages employ case particles to mark syntactic relations. In addition to fulfilling a purely grammatical function of marking the subject, the nominative markers *-ga* (Japanese) and *-i/-ka* (Korean) may bring about an additional exhaustive interpretation, illustrated in (9). Under the basic neutral description reading in (8) the subject-marked DP remains prosodically unmarked and requires no prior context. However, when combined with a prominent stress, the subject-marked DP carries an exhaustive listing reading, associated with a constituent under narrow focus (i.e., focus that does not project beyond a particular XP to any of the higher constituents in the clause). The narrow focus results in a more constrained interpretation of the marker; it signals a DP that introduces new information into discourse (e.g., an answer to a wh-question) or a similar context compatible with an exhaustive interpretation, for example the choice under 'only' or 'even', as in (10); cf. Tomioka (2009, 2010) for a detailed discussion.

(8) a. Ame-ga hutte i-masu. *Japanese*
rain-NOM falling be-PRS.DECL
'It is raining.' (Kuno 1973: 38)

b. Pi-ka nayli-ko iss-ta. *Korean*
rain-NOM falling-COMP be-PRS.DECL
'It is raining.'

(9) a. John-ga gakusei desu. *Japanese*
John-NOM student be.PRS
'It is John who is a student.' ('Of all the people under discussion, John and only John is a student')
(Kuno 1973: 38)

b. John-i haksayng i-ta. *Korean*
John-NOM student be-PRS.DECL
'It is John who is a student' ('Of all the people under discussion, John and only John is a student')

(10) (Context: Who passed the test?)

a. Hiroki-wa uka-tta. *Japanese*
Hiroki-TOP pass-PST.DECL
'(At least) Hiroki passed.'

b. Hiroki-ga uka-tta.
Hiroki-NOM pass-PST.DECL
'(Only) Hiroki passed.'

Thus, while the neutral descriptive nominative case particle fulfills a grammatical function of establishing structural relations within a sentence, its exhaustive listing counterpart interacts with the larger linguistic context and expresses distinctions related to the information structure of the utterance. Based on this observation, we will assume that the nominative case particle represents a phenomenon mediated within narrow syntax, while the exhaustive listing particle and the topic marker are linguistic elements linked to information at the level of discourse.

Table 1 presents a summary of the main functions of the Japanese and Korean particles discussed above.

**Table 1.** Functions of topic and nominative case particles in Japanese and Korean: Summary

| Particle | Functions | Particle Omissions |
|---|---|---|
| Japanese *wa*, Korean *(n)un* | 1) thematic<br>a) anaphoric<br>b) generic | yes |
| | 2) contrastive | no |
| Japanese *ga*, Korean *i/ka* | 1) nominative case | yes, in main clauses |
| | 2) exhaustive listing | no |

In terms of their syntactic position, topic particles occupy a higher structural projection than subject particles. Topics are located in TopP, housed in the CP domain, above the TP (Kishimoto, 2009; Kuroda, 1992, 2005). The nominative case particle appears in a TP-internal position (Tateishi, 1994; Ono, 2001; Ueda, 1993). This hierarchy is summarized in (11) below. We further assume that the exhaustive listing particle, associated with focus, is located in a higher syntactic position than the nominative case marker (Vermeulen, 2005), in a functional projection FocP, as shown in (12).

(11) $[_{TopP}$ DP-wa/-(n)un... $[_{TP}$ DP-ga/-i/-ka...]]

(12) $[_{FocP}$ DP-ga/-i/-ka... $[_{TP}$ DP-ga/-i/-ka...]]

Both case and topic particles may be omitted in spoken contexts and informal registers (Lee & Thompson, 1989;

Kuno, 1972; Tsutsui, 1984; Shimojo, 2006; Yatebe, 1999), as illustrated in (13) and (14):

(13) a. Kittin-ni oisii pai(-ga) ari-masu-yo. *Japanese*
kitchen-in delicious pie(-NOM) exist-PRS-PTC
'There is a delicious pie in the kitchen.'

b. Pwuek-ey masissnun pai(-ka) iss-ta. *Korean*
kitchen-in delicious pie(-NOM) be-PRS.DECL
'There is a delicious pie in the kitchen.'

(14) a. Kinoo-no-yoru tomodati-to issyo-ni mi-ta eiga(-wa) totemo yo-katta. *Japanese*
yesterday-GEN-night friends-and together-with watch-PST movie(-TOP) very be.good-PST-DECL
'The movie that I watched with my friends last night was very good.'

b. Eceysspam chingwu-wa hamkkey bo-ass-ten yenghwa(-nun) acwu coha-ess-ta. *Korean*
yesterday_night friends-and together with watch-PST-ADN movie(-TOP) very be.good-PST-DECL
'The movie that I watched with my friends last night was very good.'

However, certain restrictions occur for the particle omissions in both languages. Existing studies on particle omissions in spoken registers report that the case marker drop, restricted largely by structural constraints such as predicate type or clause type, occurs less frequently than the topic particle drop, licensed to a greater extent by pragmatic and contextual factors (Machida et al., 2004). For example, in main clauses, the nominative case marker may be omitted with stative predicates, as illustrated in example (13) above, and in unaccusative constructions, but it must always be overt in embedded clauses or when it carries the exhaustive listing interpretation (Shimojo, 2006; Takezawa, 1987; Yatabe, 1999). No omissions are possible with the contrastive topic marker, which must be overt in both languages (Yatabe, 1999); however, non-contrastive topic particles are often omitted in colloquial

speech, as illustrated by acceptable sentences in example (14) above.

## 2.2 Testing the knowledge of topics/subjects in heritage and L2 populations

The study presented below takes as its point of departure the findings of two other studies which we will briefly summarize here. A preliminary investigation of Japanese heritage speakers' production by Laleko and Kawamura (2011) showed that heritage speakers consistently underused the topic marker and overused the subject marker in retelling a story: heritage language speakers used the subject marker more frequently than the topic marker (a pattern opposite to that observed in the monolingual group) and extended the nominative case marker to anaphoric contexts where the topic marker would have been more acceptable. Furthermore, a preliminary investigation of subject and topic comprehension by Japanese and Korean heritage and L2 speakers showed that both groups were less similar to native speakers in rating sentences that involved the topic marker compared to those with the subject marker. Both groups displayed non-target-like knowledge of rules governing particle omissions (Laleko & Polinsky, 2013). Assuming that topic marking is regulated to a greater extent by pragmatic knowledge than the marking of nominative case, these findings suggest greater difficulty with discourse-level phenomena than with those mediated within the grammar. Yet, they leave us with no direct way of probing the possible causes of the observed asymmetry. Could the difficulty with topic marking be due to its status as an interface phenomenon, its structural complexity, or the memory demands necessary for evaluating the information in the topic-marked constituent against prior context?

The preliminary findings above were obtained in the context of a more general investigation into possible differences between subjects and topics in Japanese and Korean. The work reported below recognizes different types of topics and subjects, as shown in Table 1, and distinguishes three fundamental motivations for the difficulties experienced by near-native speakers: the inherent interface status of discourse-level elements; clausal

size and complexity; and the demands on short-term memory necessary for reference tracking and resolving contextual optionality. To evaluate these competing pressures, we will explore the following contrasts in Japanese and Korean:

(15) Congruous topic/subject marking

- topic
  - anaphoric
  - generic
  - contrastive
- subject
  - descriptive
  - exhaustive

(16) Incongruous marking

a. topic marker used in place of subject marker, for all the subtypes in (15a)

b. subject marker used in place of topic marker, for all the subtypes in (15b)

(17) Omission of topic marker

a. congruous

b. incongruous

(18) Omission of subject marker

a. congruous

b. incongruous

Examples of test sentences that reflect the desiderata in (15)-(18) are presented in the Appendix. In this article, we report the results of conditions (15) and (17)-(18) only. To anticipate the discussion below, we will not consider more fine-grained differences between Japanese and Korean with respect to topic/subject marking (but see Shimojo & Choi, 2000). These differences are important, but since we do not have a full range of data on the more basic contrasts illustrated in (15)-(18), such contrasts should be established first.

## 2.3 Hypotheses and predictions

Recall that our umbrella hypothesis posits that discourse-level phenomena incur greater difficulty than phenomena contained within narrow syntax; therefore, we expect heritage speakers and L2 learners to exhibit better knowledge of the subject marker than the topic marker. In the context of an experiment involving acceptability ratings, this pattern could be observed if the marker that presents more difficulty for the bilingual populations receives ratings that diverge to a greater extent from those of the monolingual baseline speakers. Conversely, ratings for the relatively easier marker could be closer to those of the monolingual controls. In this scenario, sentences using the topic marker inappropriately could be overrated by the bilingual participants to a greater extent than sentences with the incongruous subject marker, and the opposite asymmetry could emerge for grammatical sentences.

Our further predictions stem from the differentiated functions of the two sets of markers, which may provide additional insights into the nature of the observed difficulties by examining the three possible explanations for the asymmetry between the syntax and discourse components of language. We will refer to these explanations as the Integration Difficulty Hypothesis, the Structural Complexity Hypothesis, and the Contextual Embedding Hypothesis. In evaluating the predictions of these three hypotheses relative to our data, we will examine the speakers' ratings of sentences containing the anaphoric, generic, and contrastive topic markers, including contexts in which these markers are overt and contexts in which they are omitted, in order to try to tease apart the underlying factors

potentially contributing to the relatively greater difficulty associated with topic-marking.

Comparing anaphoric, generic, and contrastive topics allows us to gain new insights into the nature of topics in general. Recall that, within the thematic topic category, anaphoric topics establish reference relations in discourse and are linked to prior discourse through an antecedent, while generic topics, which also mark the theme of the utterance, are not dependent upon prior discourse in the same way. In other words, the referents of generic themes occur outside the discourse, and are in that sense discourse-external, whereas anaphoric themes, operating discourse-internally, are always embedded into prior context (Iwasaki, 1987); see also Halliday and Hasan's (1976) "homophoric" and "endophoric" distinction. According to the Contextual Embedding Hypothesis, elements that require referential linking to other material in prior discourse are associated with greater difficulty than elements that do not require such linking for their construal and interpretation. The Contextual Embedding Hypothesis thus predicts differentiated results in anaphoric and generic conditions. Discourse-internal theme maintenance and tracking through a discourse-internal topic particle should be more costly in terms of attention and memory resources than theme creation, associated with a generic topic marker. If this prediction is on the right track, we can expect decreased accuracy in the anaphoric topic condition compared to the non-anaphoric generic condition.

However, under the generalized Integration Difficulty Hypothesis, which attributes the difficulty of a particular phenomenon to the need to coordinate knowledge from different domains, such as syntax and pragmatics, we should expect no differences in accuracy between anaphoric and generic topics. Both types of thematic topics mark pragmatically-relevant distinctions related to the expression of old and new information in discourse: namely, they mark the theme of the utterance. Hence, both types of thematic topics are presumably mediated within the syntax-pragmatics interface, also known as the external interface, which integrates linguistic information with contextual information external to the grammar (Sorace & Serratrice 2009; Sorace, 2011).

The Structural Complexity Hypothesis, which takes the hierarchical syntactic representation as its point of departure, makes a similar prediction: Under the assumption that all thematic elements are contained within the same syntactic projection, TopicP, no differences are expected between the anaphoric and generic conditions. However, the same hypothesis predicts a difference between thematic and contrastive topics, categories associated with two separate syntactic projections. If thematic topics are located in the highest position in the tree structure (Heycock, 1998: 77) and higher than contrastive topics, they should involve more structure than contrastive topics. Thus, the Structural Complexity Hypothesis predicts greater difficulty with thematic topics, compared to contrastive topics.

In addition to examining the three subcategories of topics, we will also analyze the descriptive and exhaustive subjects in the two languages. Comparing the two functions of the nominative case marker will grant additional insight into the general syntax-before-discourse issue. If the umbrella hypothesis is correct, the neutral description reading associated with the core nominative case function will be easier than the narrow focus reading, which interacts to a greater extent with the information structure and the larger linguistic and non-linguistic context of the utterance. However, this pattern would be compatible with all three accounts, which in this instance yield similar predictions. The exhaustive reading marker represents an interface phenomenon (whereas the nominative case marker is mediated entirely within syntax), occupies a higher syntactic position, FocP (while the nominative marker is housed within the lower TP), and, due to its link with prior discourse, places a greater burden on memory than the marker of structural nominative case.

The hypotheses and predictions are summarized in Table 2 below.

**Table 2.** Hypotheses and predictions

| | Umbrella Hypothesis | Integration Difficulty Hypothesis | Structural Complexity Hypothesis | Contextual Embedding Hypothesis |
|---|---|---|---|---|
| | topic > subject | | | |
| Topic | | anaphoric = generic = contrastive | thematic > contrastive | anaphoric > generic |
| Subject | | exhaustive > descriptive | exhaustive > descriptive | exhaustive > descriptive |

= : equal ease of processing; > : more difficult to process

An additional question that we will address in this study concerns the nature of linguistic deficits under incomplete language acquisition. If heritage or L2 speakers exhibit non-target-like knowledge of the topic or case particles, how should these results be interpreted? In other words, do these speakers experience problems with the *semanticrepresentation* of these categories, or are the difficulties primarily with surface morphological *marking*? [2] In order to address this question, we examine conditions where the case and topic particles are overtly expressed and contexts where they are omitted. If heritage speakers and second language learners exhibit unequal difficulty in contexts that involve zero marking and contexts that require overt particles, this would point to morphology (or mapping between morphology and internal representation) as the likely source of difficulty. Under the assumption that semantic representations are the same in null and overt contexts, equal difficulty in these conditions could signal representational problems.

## 3. Participants and methodology

64 heritage language speakers (*N* = 29 for Japanese; *N* = 35 for Korean) and 47 second language learners (*N* = 31 for Japanese; *N* = 16 for Korean) participated in the study. All heritage language speakers were exposed to the heritage language at home while growing up, but reported another language as their current main language of communication. All second language learners had an average of 3 years of formal instruction in Japanese or Korean and reported not hearing or speaking the second language at home or in another naturalistic setting. The control group for each language included 15 age-matched native speakers. Table 3 below summarizes the relevant mean demographic information for the target groups of heritage language (HL) speakers and second language (L2) learners.

**Table 3.** Participants: Demographic Information

| Language | Korean | | Japanese | |
|---|---|---|---|---|
| Group | L2 (*N* = 16) | HL (*N* = 35) | L2 (*N* = 31) | HL (*N* = 29) |
| Age | 25.8 | 24.5 | 27.5 | 24.7 |
| Age of departure from Korea/Japan | N/A | 3.9 | N/A | 4.5 |
| Age of switch to another language | N/A | 3.9 | N/A | 5.3 |
| Daily use of Korean/Japanese (%) | 23.5 | 32.9 | 12.4 | 22.9 |
| Self-rated proficiency in Korean/Japanese (1–5) | 3.39 | 4.35 | 3.01 | 3.62 |

All bilingual participants were asked to provide detailed self-ratings of their proficiency in the non-dominant

language (Korean or Japanese) across the four domains of language competence: listening, reading, speaking, and writing. The ratings were obtained using a scale ranging from 1 to 5, where 5 was the highest possible rating. Table 4 below presents the percent of responses for each possible rating in each language group.

**Table 4.** Participants: Self-Rated Language Proficiency (1–5 scale)

| | Korean | 1 | 2 | 3 | 4 | 5 | Japanese | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Listening | HL | 7 | 9 | 14 | 19 | 51 | HL | 5 | 5 | 25 | 40 | 25 |
| | L2 | 0 | 18 | 29 | 24 | 29 | L2 | 11 | 14 | 36 | 32 | 7 |
| Reading | HL | 12 | 12 | 19 | 12 | 45 | HL | 10 | 10 | 35 | 25 | 20 |
| | L2 | 0 | 26 | 31 | 12 | 31 | L2 | 7 | 21 | 39 | 21 | 12 |
| Speaking | HL | 12 | 11 | 9 | 19 | 49 | HL | 10 | 10 | 20 | 40 | 20 |
| | L2 | 6 | 19 | 25 | 25 | 25 | L2 | 11 | 25 | 32 | 18 | 14 |
| Writing | HL | 22 | 6 | 20 | 12 | 40 | HL | 20 | 15 | 20 | 25 | 20 |
| | L2 | 19 | 19 | 19 | 7 | 36 | L2 | 11 | 18 | 50 | 21 | 0 |

In the main experiment, the participants were asked to provide acceptability ratings for 56 sentences, using a 1–5 scale. The experimental conditions followed the contrasts in (15)-(18) above and included congruous use of the topic and subject markers, misuse of the markers, and congruous/incongruous particle omissions. All the sentences were presented in written form, which means that all the participants in the experiment had to be literate in Korean or Japanese. This condition may exclude some heritage speakers, who often lack literacy in their heritage language (Benmamoun, Montrul, & Polinsky, 2013b). Examples of the test sentences appear in the appendix.[3] The ratings were elicited using Amazon Mechanical Turk (Gibson, Piantadosi, & Fedorenko, 2011;

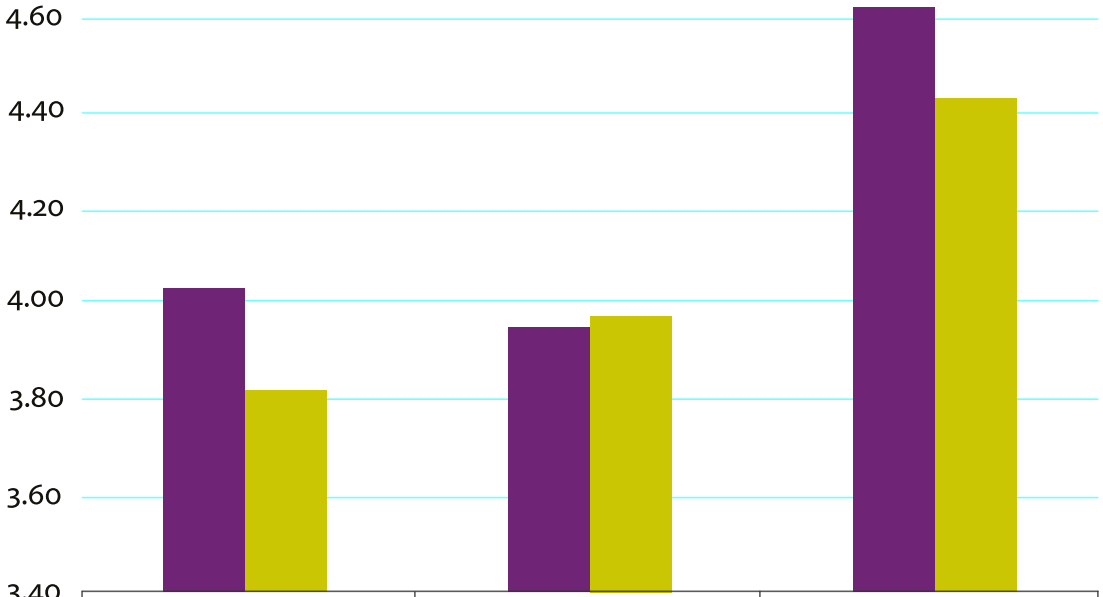
Sprouse, 2011). The study also involved a detailed sociolinguistic background questionnaire, which all participants completed in their dominant language.

## 4. Results and discussion

### 4.1 Overview of results: Topic and subject particles

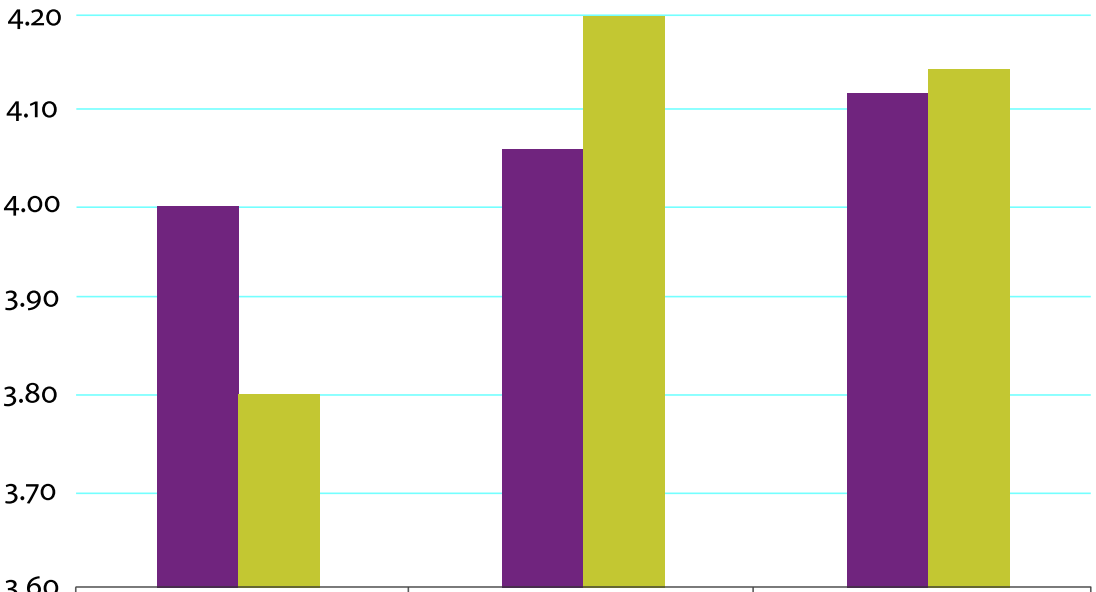
Two-sample t-tests with a Bonferroni adjustment were performed to evaluate differences between mean ratings for the contrasts tested in the study. Mean ratings for sentences involving appropriate uses of the topic particle (TOP) and the descriptive subject particle (we will refer to the non-exhaustive subject particle as the nominative case (NOM) marker) are presented in fig 1 for Japanese and in fig 2 for Korean. In each language group, heritage language speakers exhibited a significant difference between these two conditions ($p < 0.01$), with higher ratings obtained in the nominative condition than in the topic condition. Speakers in all remaining groups exhibited no statistical differences between the NOM and TOP conditions with overt appropriately used particles.

**Figure 1.** Mean ratings for overt nominative case (NOM) and topic (TOP) particles used in appropriate contexts: heritage speakers (HJ), L2 learners (L2J), and native speakers (L1J) of Japanese




| | HJ | L2J | L1J |
|---|---|---|---|
| NOM | 4.03 | 3.94 | 4.62 |
| TOP | 3.81 | 3.96 | 4.46 |

**Figure 2.** Mean ratings for overt nominative case (NOM) and topic (TOP) particles used in appropriate contexts: heritage speakers (HK), L2 learners (L2K), and native speakers (L1K) of Korean

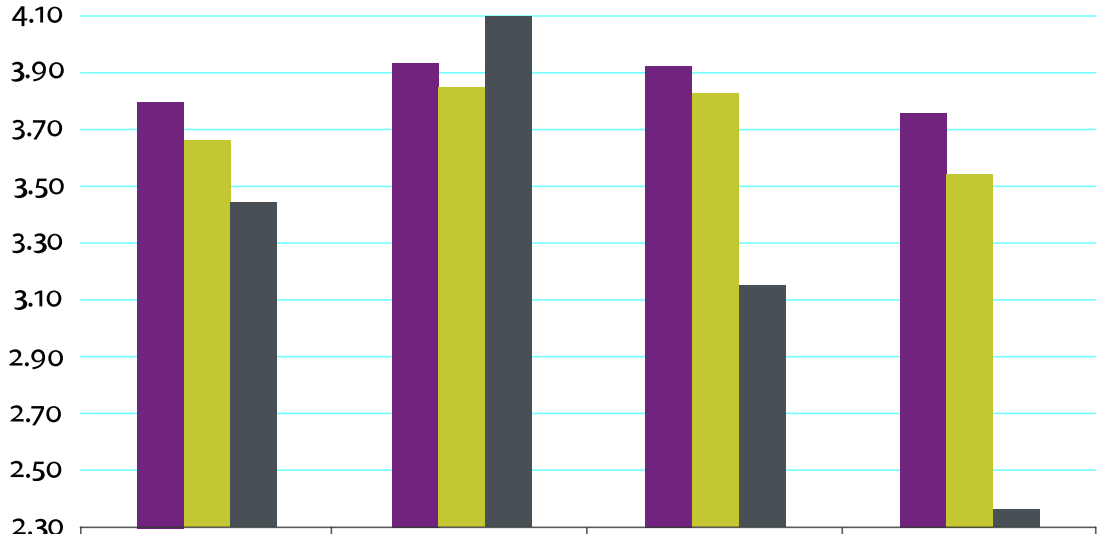



| | HK | L2K | L1K |
|---|---|---|---|
| NOM | 4.00 | 4.06 | 4.12 |
| TOP | 3.80 | 4.24 | 4.14 |

Next, we present the results for conditions involving particle omissions. fig 3 and 4 below summarize the mean ratings for the congruous and incongruous omissions of topic and subject particles in the two languages under investigation. All groups of participants with the exception of second language learners of Korean exhibited a significant contrast ($p < 0.01$) between acceptable and unacceptable omissions of the subject marker. This finding suggests that most bilingual speakers in our study were able to make a distinction between contexts in which omissions of the subject marker are permissible and contexts in which such omissions result in ungrammaticality.

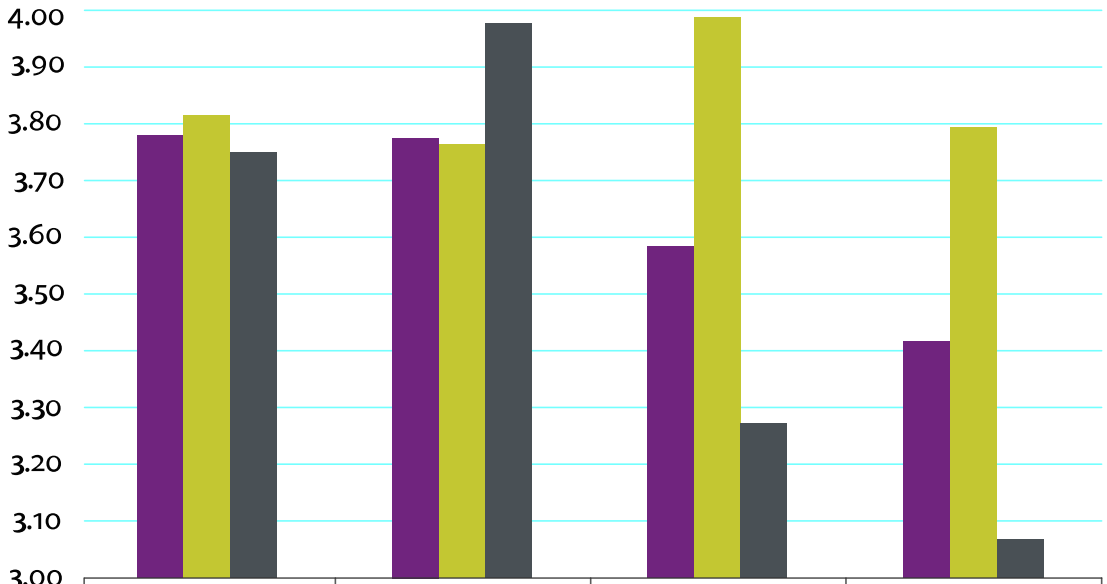
In contrast, only monolingual speakers showed a reliably significant contrast between acceptable and unacceptable omissions of the topic particle ($p < 0.01$). All bilingual groups performed at chance when rating particle omissions in topic. Recall that omissions of the subject particle are constrained predominantly by syntactic factors, while omissions of the topic particle in both languages occur more freely and may be deemed as appropriate or inappropriate relative to the specific contextual factors, level of formality, and speaker's intent. Lack of clear judgments in heritage speakers and second language learners of Japanese and Korean with respect to where topic particles may or may not be omitted lends support to the idea that constructions involving pragmatic and discourse-level knowledge may be more difficult to interpret for these speakers than constructions regulated by the grammar.

**Figure 3.** Particle omissions: mean ratings for congruously and incongruously (#) omitted topic (TOP) and subject (SUBJ) particles: heritage speakers (HJ), L2 learners (L2J), and native speakers (L1J) of Japanese




| | (TOP) | (SUBJ) | #(TOP) | #(SUBJ) |
|---|---|---|---|---|
| HJ | 3.78 | 3.97 | 3.93 | 3.78 |
| L2J | 3.66 | 3.83 | 3.82 | 3.55 |
| L1J | 3.44 | 4.11 | 3.14 | 2.37 |

**Figure 4.** Particle omissions: mean ratings for congruously and incongruously (#) omitted topic (TOP) and subject (SUBJ) particles: heritage speakers (HK), L2 learners (L2K), and native speakers (L1K) of Korean




| | (TOP) | (SUBJ) | #(TOP) | #(SUBJ) |
|---|---|---|---|---|
| HK | 3.78 | 3.76 | 3.58 | 3.41 |
| L2K | 3.83 | 3.75 | 3.99 | 3.79 |
| L1K | 3.73 | 3.97 | 3.26 | 3.07 |

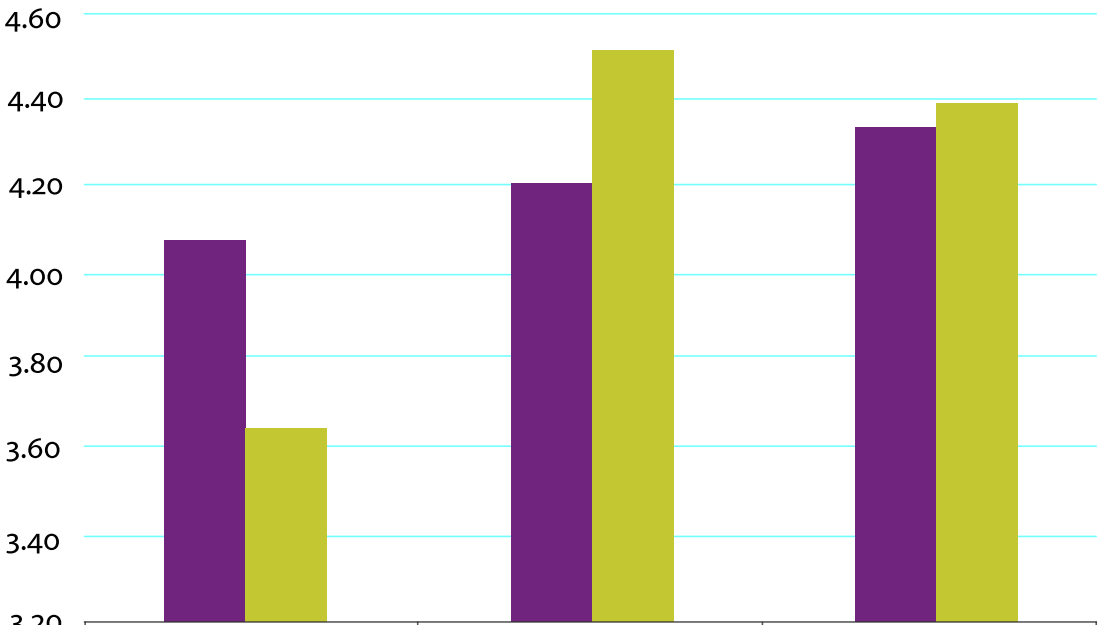
Ratings obtained from the monolingual speakers offer additional empirical evidence for the previously made observation that different types of constraints are likely to be involved in licensing the omissions of topic particles and nominative case particles in the two topic-prominent languages under investigation. Unacceptable topic particle omissions were rated consistently higher than incongruous omissions of the subject particle, whereas the opposite pattern was observed for acceptable omissions of the topic and subject particles, in which higher ratings

were obtained for omitted subjects. In other words, a considerably wider gap was observed between acceptable and unacceptable omissions of the subject marker, compared to the relatively smaller gap between felicitous and infelicitous omissions of the topic marker. Topic particle omissions in both languages were associated with more gradient judgments in native speakers, while omissions of the subject particle triggered more categorical judgments. These patterns are reflective of the previously documented distinction between soft and hard constraints, proposed to account for a certain degree of variability in judgments associated with some linguistic phenomena in monolingual speakers (as well as optionality and variability in bilinguals). Soft constraints are to a large degree context-dependent and typically trigger gradient judgments in monolinguals, with violations resulting in mild unacceptability. In contrast, hard constraints are those where native speakers' judgments tend to be more categorical (Sorace & Keller, 2005). Ratings on particle omissions obtained from the monolingual speakers reflect this distinction: omissions of the nominative case particle revealed a pattern typical of hard constraints, whereas null topic particles yielded results consistent with those observed for soft constraints. As evidenced by a uniform treatment of pragmatically felicitous and infelicitous topic particle omissions in heritage speakers and second language learners of Japanese and Korean, the bilingual speakers in this study were not sensitive to differences in acceptability governed by soft constraints, an area where developmental optionality and instability is predicted for these populations (Sorace & Keller, 2005). At the same time, the majority of bilingual speakers remained target-like in recognizing the contrast between grammatical and ungrammatical omissions of the subject particle, an area where hard constraints appear to be involved.

Next, we examine the ratings for the two separate functions of the subject marker, descriptive and exhaustive. Since each function is associated with distinct interpretive and syntactic properties, as discussed earlier, a difference in acceptability judgments for these separate functions could signal unequal difficulty in interpreting these markers in some populations of speakers. If such difference is detected, greater difficulty should be

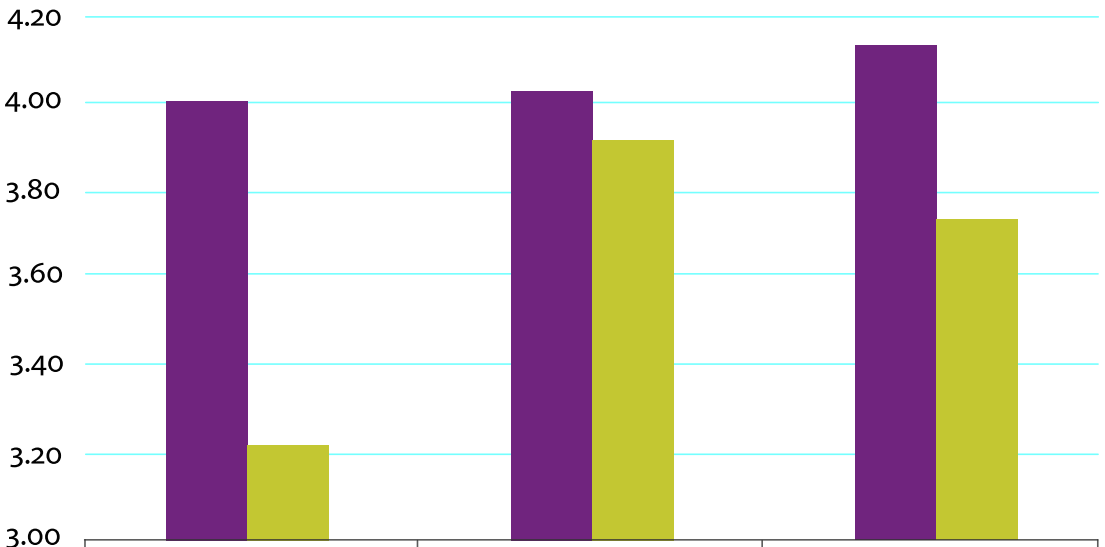
expected for the exhaustive subject, linked with the information structure of the utterance through its association with narrow focus. This prediction was born out for heritage language speakers of Japanese and Korean. Sentences with the descriptive subject marker, associated with the structural nominative case, received significantly higher ratings in the groups of heritage speakers in both languages than sentences with the exhaustive listing particle ($p < 0.01$ in Japanese and $p < 0.001$ in Korean). No other groups showed significant differences between these conditions. These results are presented in fig 5 and fig 6 below.

**Figure 5.** Descriptive and exhaustive subject particles: heritage speakers (HJ), L2 learners (L2J), and native speakers (L1J) of Japanese




| | HJ | L2J | L1J |
|---|---|---|---|
| descriptive | 4.03 | 3.94 | 4.62 |
| exhaustive | 3.65 | 4.00 | 4.46 |

**Figure 6.** Descriptive and exhaustive subject particles: heritage speakers (HK), L2 learners (L2K), and native speakers (L1K) of Korean




| | HK | L2K | L1K |
|---|---|---|---|
| descriptive | 4.00 | 4.06 | 4.12 |
| exhaustive | 3.24 | 3.91 | 3.75 |

The exhaustive marker, whose appropriate use requires keeping track of contextual information, is more difficult than the sentence-level descriptive nominative particle, mediated in narrow syntax and not linked explicitly to larger discourse. This result corroborates the overall pattern predicted by our umbrella hypothesis (i.e., that discourse computation presents more challenges than syntactic computation) for heritage language speakers, but it leaves us with no direct way of probing into the underlying factors in the context where all three accounts converge on their predictions. In the following section, we turn to the specific hypotheses that may guide us toward a possible

explanation of this apparent asymmetry.

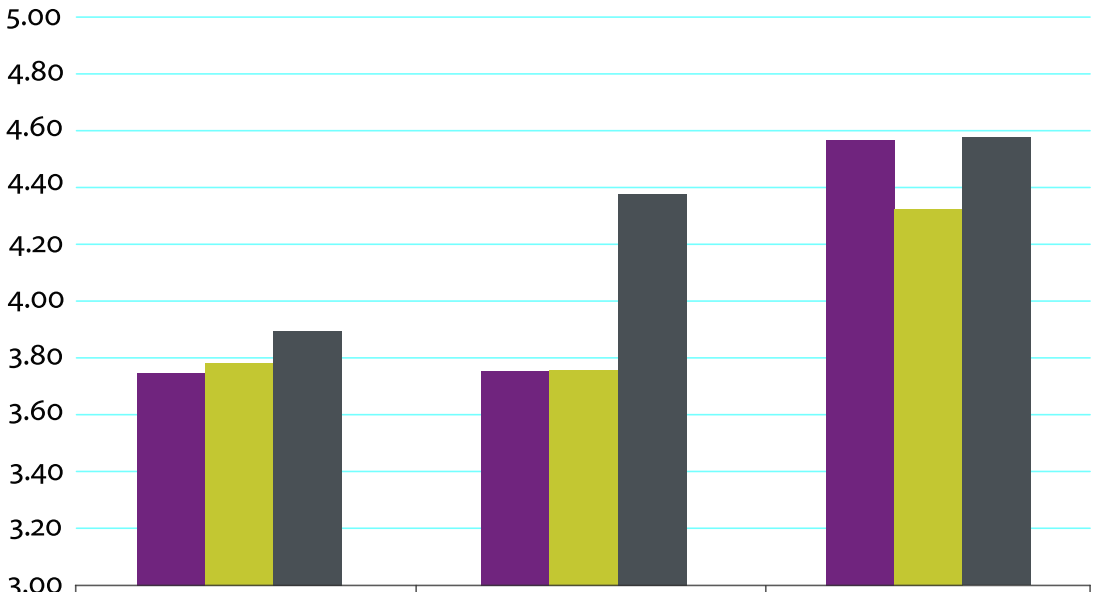
## 4.2 Anaphoric, generic, and contrastive topics

To summarize the overall results discussed so far, we observe that heritage speakers and L2 learners of Japanese and Korean exhibit some difficulty in contexts involving the topic marker, but are generally more on target in contexts that involve the subject marker. This pattern occurs in differences observed in conditions involving overt as well as null particles. In both types of conditions, bilingual speakers consistently exhibit more uncertainty in rating sentences targeting their knowledge of the topic marker. In contexts involving particle omissions, even native speakers exhibited different patterns with respect to their ratings of topic and subject particle omissions: topic marker omissions had more gradient acceptability than subject marker omissions. Contrasts between acceptable and unacceptable topic particle omissions (an area governed by soft constraints, as discussed above) were altogether undetectable in the data from the bilingual speakers, who were overall more successful in differentiating between grammatical and ungrammatical omissions of the subject particle. An additional contrast between ratings for the two separate functions of the subject particle, descriptive and exhaustive, shown by heritage speakers of both languages under investigation, lends further support to the general observation that whenever a difference between syntax-level and discourse-level elements appears in bilingual speakers, discourse-level elements represent a greater challenge.

Keeping these results in mind, we now turn to the main question of our article: to compare three specific hypotheses that may provide further insight into the nature of the observed asymmetry between syntactic and discourse knowledge in bilingual grammars. To do so, we will examine the ratings for the three separate functions of the topic marker (contrastive, anaphoric, and generic) in different conditions, relative to the predictions formulated in Section 2.3 above.

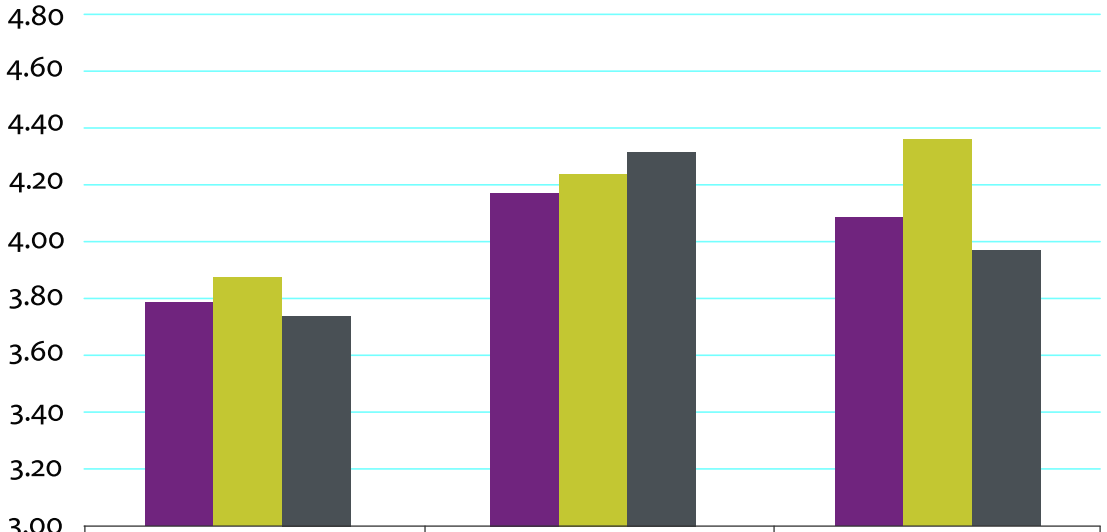
First, we present the mean ratings for overtly used topic particles in generic, anaphoric, and contrastive contexts for the three groups of speakers of Japanese (fig 7) and Korean (fig 8).

**Figure 7.** Overtly used topic particles in generic, anaphoric, and contrastive contexts: heritage speakers (HJ), L2 learners (L2J), and native speakers (L1J) of Japanese




| | HJ | L2J | L1J |
|---|---|---|---|
| contrastive | 3.75 | 3.76 | 4.54 |
| anaphoric | 3.78 | 3.76 | 4.27 |
| generic | 3.90 | 4.37 | 4.58 |

**Figure 8.** Overtly used topic particles in generic, anaphoric, and contrastive contexts: heritage speakers (HK), L2 learners (L2K), and native speakers (L1K) of Korean



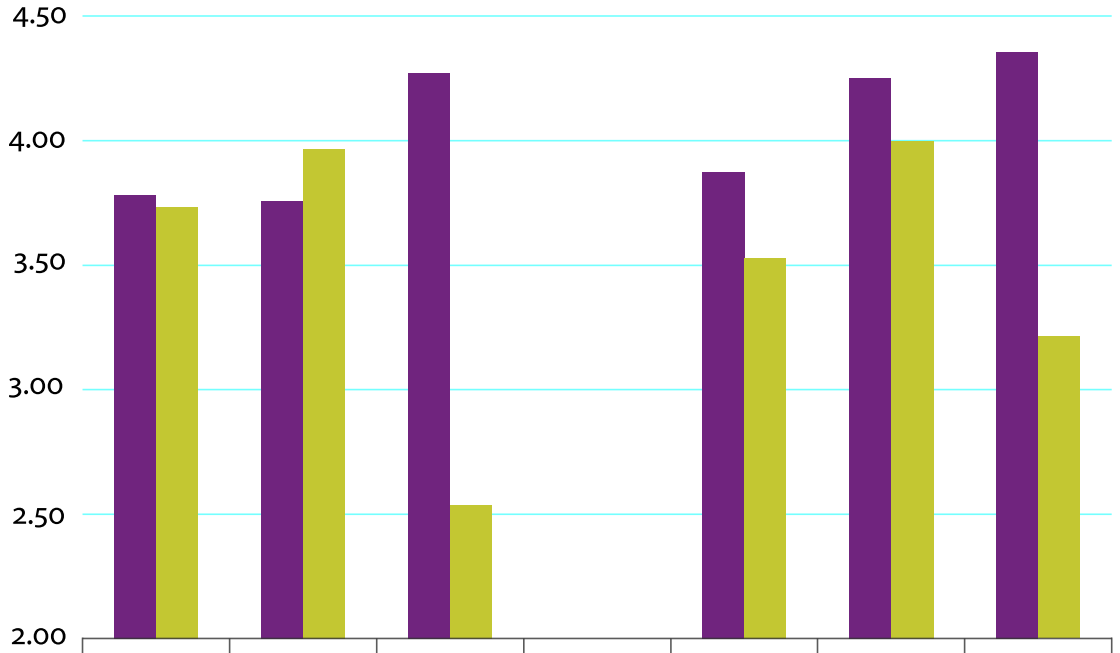

| | HK | L2K | L1K |
|---|---|---|---|
| contrastive | 3.79 | 4.16 | 4.11 |
| anaphoric | 3.88 | 4.25 | 4.36 |
| generic | 3.74 | 4.31 | 3.96 |

In the overt particle condition, the three functions of the topic particle were treated uniformly by most groups, with the exception of second language learners of Japanese. In this group, two statistically significant differences appeared, both involving generic topics. Sentences with generic topics were rated significantly higher than sentences with anaphoric topics ($p < 0.001$) and sentences with contrastive topics ($p < 0.001$). These findings suggest that whenever differences among the three functions of the topic marker are apparent in the ratings of acceptable sentences containing these particles, the generic interpretation of the topic marker appears to be

relatively easier to obtain than its contrastive or anaphoric readings.

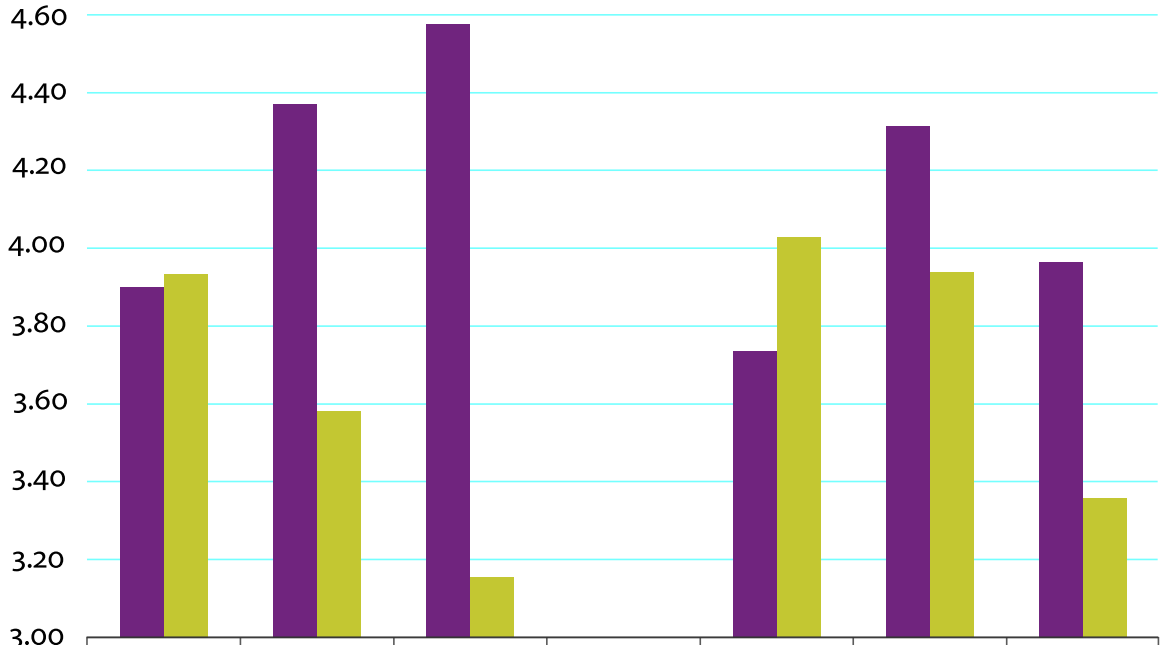
Next, we examine incongruous omissions of the topic marker in its anaphoric (fig 9), generic (fig 10), and contrastive (fig 11) contexts. In all of these contexts, an overt marker would have been the more acceptable choice. Mean ratings for the incongruous omissions are presented below. Mean ratings for sentences with overtly used particles in each of the conditions are also included for comparison.

**Figure 9.** Mean ratings for overt and incongruously omitted anaphoric topic particles for heritage speakers, second language learners, and native speakers of Japanese and Korean



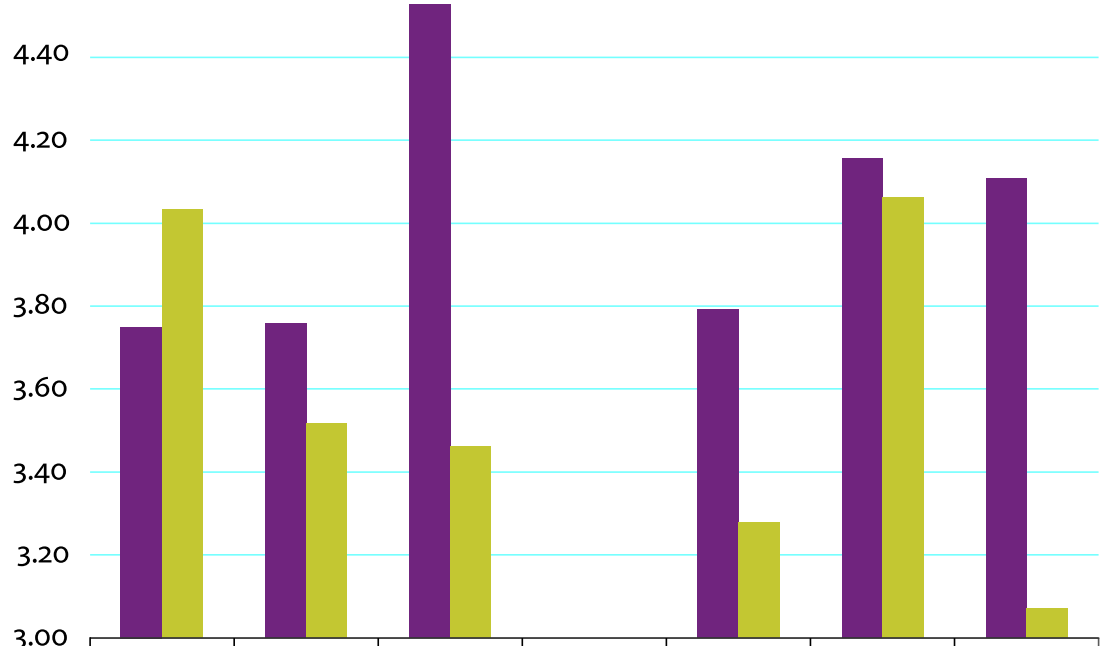

| | HJ | L2J | L1J | | HK | L2K | L1K |
|---|---|---|---|---|---|---|---|
| anaphoric overt | 3.78 | 3.76 | 4.27 | | 3.88 | 4.25 | 4.36 |
| #anaphoric null | 3.73 | 3.97 | 2.54 | | 3.53 | 4.00 | 3.21 |

**Figure 10.** Mean ratings for overt and incongruously omitted generic topic particles for heritage speakers, second language learners, and native speakers of Japanese and Korean




| | HJ | L2J | L1J | | HK | L2K | L1K |
|---|---|---|---|---|---|---|---|
| generic overt | 3.90 | 4.37 | 4.58 | | 3.74 | 4.31 | 3.96 |
| #generic null | 3.93 | 3.58 | 3.15 | | 4.03 | 4.94 | 3.36 |

**Figure 11.** Mean ratings for overt and incongruously omitted contrastive topic particles for heritage speakers, second language learners, and native speakers of Japanese and Korean




| | HJ | L2J | L1J | | HK | L2K | L1K |
|---|---|---|---|---|---|---|---|
| contrastive overt | 3.75 | 3.76 | 4.54 | | 3.79 | 4.16 | 4.11 |
| contrastive overt | 4.03 | 3.52 | 3.46 | | 3.28 | 4.06 | 3.07 |

In all three functions of the topic marker (anaphoric, generic, and contrastive), monolingual speakers exhibited significant differences ($p < 0.05$) between conditions in which the topic particle was used overtly and conditions in which it was omitted, resulting in unacceptability. However, within-group comparisons performed for each group of the Japanese and Korean bilingual speakers in the anaphoric, generic, and contrastive conditions revealed that heritage language speakers and second language learners of Japanese and Korean did not exhibit target-like

acceptability contrasts in the anaphoric contexts. Differences between conditions involving overt and infelicitously omitted anaphoric topic particles were not significant in either bilingual group, in either language ($p > 0.05$). In contrast, target-like acceptability patterns were observed in some bilingual populations for generic and contrastive markers. Second language learners of Japanese exhibited a reliable statistical contrast ($p < 0.001$) between acceptable overt and unacceptable null markers of generic topics. Similarly, heritage speakers of Korean distinguished ($p = 0.05$) between conditions involving overt and incongruously omitted contrastive topic particles. These findings indicate that while no bilingual group exhibited target-like knowledge of omissions in the anaphoric condition, some populations were nevertheless successful in recognizing unacceptability in generic and contrastive contexts.

In order to obtain a more detailed picture of how our bilingual speakers compared to the monolingual controls on their knowledge of the specific functions of the topic markers, we further ran a series of across-group comparisons for conditions involving unacceptable omissions of generic, anaphoric, and contrastive topic particles. Once again, the generic topic marker seemed to be the least problematic topic particle in all populations. Both heritage language speakers and second language learners of Japanese were not statistically distinguishable from the native Japanese speakers on their ratings of topic particle omissions in the generic condition ($p > 0.05$). Similarly, both heritage speakers and second language learners of Korean patterned with the speakers in the monolingual Korean group in rating sentences with incongruous omissions of the generic topic marker ($p > 0.05$). However, differences were observed in the anaphoric topic condition between heritage speakers of Japanese and monolingual controls ($p < 0.01$) and between second language learners of Japanese and monolingual controls ($p < 0.001$). Additionally, second language learners of Korean differed from monolingual controls on particle omissions in the contrastive topic condition ($p < 0.05$). Thus, sentences testing the anaphoric function of the topic marker as well as sentences testing its contrastive function were rated in a non-target-like way by some groups of

bilinguals, but generic topics were never problematic for any of the four bilingual populations in the two languages under investigation.

### 4.3 Discussion

The results of our study corroborate the observation that information structure-level phenomena present relatively greater challenges for bilingual speakers than those mediated within syntax. Sentences involving topic particles were generally more problematic for both early and late bilinguals than sentences targeting the knowledge of the subject marker, although some late bilinguals exhibited problems with both markers when rating omissions, possibly a reflection of a more global difficulty with ellipsis. Additionally, heritage language speakers in both language groups differentiated between two separate functions of the subject marker. The descriptive subject, a marker of grammatical case at the sentential level, was easier for these speakers to evaluate than the exhaustive listing subject, associated with new information at the discourse level.

These results could be accounted for with reference to the Interface Hypothesis. By virtue of their contribution to the information structural partition of the utterance, topics represent the syntax-pragmatics interface and require both syntactic and discourse-level knowledge, whereas grammatical case represents a syntactic phenomenon that is not dependent on discourse. Greater problems with topic marking could then be a consequence of an extra processing load imposed by the need to coordinate information from multiple domains (cf. Koornneef, 2008 for similar considerations with respect to the processing load imposed by domain integration).

At the same time, these results could also follow from a structural asymmetry between topics and subjects. Topics, which occupy a syntactic position in the left periphery, are located higher in the syntactic structure than subjects, generated below the CP level. Assuming that syntactic projections are built from the bottom up, the

presence of a higher projection entails the presence of all lower projections contained therein. If building more syntactic structure requires more computational effort, greater difficulty with topics is predicted without reference to the integration problem.

Differential results in conditions involving three specific functions of the topic marker (anaphoric, generic, and contrastive) observed in some bilingual groups allowed us to examine the predictions of the three hypotheses in more detail. Such differential results for the distinct functions of topics in Japanese and Korean are unexpected under the generalized interface-based account offered by Integration Difficulty Hypothesis, which distinguishes interface domains from non-interface domains, as well as external interfaces from internal interfaces, but does not differentiate phenomena mediated within the same interface. Assuming that all subtypes of topics represent the same external interface (i.e., the syntax-pragmatics interface), the integration problem alone cannot solely account for any observed asymmetry among them. Thus, it appears that the interface status may not sufficiently explain varying degrees of difficulty associated with the specific functions of the topic marker.

The analysis of the mean ratings for acceptable sentences with overt topic particles revealed that some bilingual groups consistently assigned higher ratings to sentences with generic topics, compared to sentences with anaphoric and contrastive topics. Assuming that acceptability judgments are correlated (inversely) with processing complexity (Fanselow & Frisch, 2006), these findings may suggest that anaphoric and contrastive topics may be more difficult to evaluate for some bilingual populations than generic topics. This difference between anaphoric and generic conditions is consistent with the predictions of the Contextual Embedding Hypothesis, which assigns greater difficulty to elements whose successful interpretation requires linking to some other material in prior context. The anaphoric topic marker, which refers back to previous discourse, should thus be more difficult than the marker of theme that occurs in non-anaphoric generic statements. This outcome is expected assuming that the acceptance of a referent (theme-creation) is less costly in processing terms than

referent maintenance, which requires access to a longer portion of discourse.

Ratings for incongruous particle omissions also shed light on the distinct functions of the topic marker. All the bilingual groups in our study had difficulty with sentences involving anaphoric topics: neither group of bilinguals exhibited target-like acceptability contrasts in the anaphoric condition. At the same time, some bilingual groups successfully recognized differences between overt and incongruously omitted particles in generic and contrastive conditions. We take these results as evidence that rules governing particle omissions with generic and contrastive topics are in principle acquirable by bilingual speakers. Rules for particle omission in the anaphoric domain, however, do not appear to have been fully acquired in any of the bilingual groups in our study. Across-the-board problems with anaphoric topics, alongside partial success in the generic condition, reveal an asymmetry between anaphoric and generic topics that lends support to the Contextual Embedding Hypothesis, which correctly predicts relatively more effort in evaluating anaphoric elements in theme due to their referential nature and connection to prior discourse. The other asymmetry, one that involves anaphoric and contrastive conditions, follows from predictions of the Structural Complexity Hypothesis, which links greater complexity of linguistic phenomena to the higher syntactic position and the more elaborate structure of syntactic projections housing them. As we discussed earlier, researchers have proposed that anaphoric topics, a subtype of thematic topics, occupy a higher structural position than contrastive topics (Heycock, 2008). Assuming such a structure, the greater difficulty with anaphoric topics, compared to contrastive topics housed in a lower projection, may stem from the additional computation efforts associated with building more structure.

Overall, the Contextual Embedding Hypothesis has greater explanatory power than its alternatives in accounting for the observed patterns in the data. The Structural Complexity Hypothesis correctly predicts the various degrees of difficulty observed for anaphoric and contrastive topics. However, assuming that generic topics are represented in the same projection as anaphoric topics, the structural explanation does not account for the

apparent asymmetry between these two types of thematic topics.

Finally, we would like to turn to the question of whether the non-target-like knowledge exhibited by the bilingual speakers in our experiment stems from representational difficulties or is more likely a consequence of other problems. The differential rates of difficulty on conditions involving overt and omitted particles demonstrated by heritage speakers and L2 learners suggest that the comprehension problems they display likely represent morphological or pragmatic deficits rather than representational deficits; they lag behind the native controls in recovering the missing information, as evidenced by non-target-like patterns of ratings observed in conditions involving null particles, but have relatively fewer problems interpreting sentences with overt markers. We would like to suggest that increased difficulty with null forms stems from their amplified ambiguity, compared to elements that are overtly marked and therefore less ambiguous. In order to choose and reinstate the most appropriate candidate, the speaker needs to evaluate a variety of options contextually available for a given expression. Topic particle omissions may therefore be particularly hard since their evaluation may require considering not only the grammatical material contained within the clause itself, but also the larger extra-sentential context.

## 5. Conclusions

This study examined phenomena mediated at different levels of linguistic organization in heritage language speakers and second language learners of Japanese and Korean. Overall, phenomena at the level of information structure were associated with greater difficulty for bilinguals than those contained within syntax. This pattern was manifested in two areas. First, bilingual speakers generally exhibited more problems in conditions that targeted their knowledge of the topic marker, compared to sentences involving the nominative case marker. Second, a difference emerged between descriptive and exhaustive functions of the subject marker. For all significant contrasts detected, the more difficult condition was the one associated to a greater extent with discourse-

pragmatic, rather than purely syntactic, knowledge. This trend is consistent with the claim that formal features associated with the syntax component of language present fewer challenges for bilingual populations than discursive features, which also involve pragmatic knowledge, and that speakers may remain non-target-like with respect to the latter group of phenomena while converging with the baseline controls with respect to the former properties.

Further comparisons of the distinct functions of the topic markers in both languages pointed to differentiated knowledge of these functions among the bilingual speakers, as evidenced by contrasts between anaphoric and generic topics, as well as by contrasts between anaphoric and contrastive topics. The asymmetry between these separate functions of the topic marker makes the explanation in terms of a generalized interface-related deficit untenable; instead, a finer analysis of the factors contributing to the interface problem is needed. Here we explored two such alternative accounts, one appealing directly to the syntactic structure of the discussed phenomena, the other linked to their referential properties. Both accounts fared better than the generalized interface-based account in predicting contrasts between different types of topics, but the approach based on the degree of contextual dependence had the most potential to explain the observed differences. These results are summarized in Table 5 below.

**Table 5.** Summary of findings

| | Umbrella Hypothesis | Integration Difficulty Hypothesis | Structural Complexity Hypothesis | Contextual Embedding Hypothesis |
|---|---|---|---|---|
| | topic > subject<br>YES | | | |
| Topic | | anaphoric =<br>generic =<br>contrastive<br>not supported | thematic >
contrastive<br>partially supported | anaphoric >
generic<br>supported |
| Subject | | exhaustive >
descriptive<br>supported | exhaustive >
descriptive<br>supported | exhaustive >
descriptive<br>supported |

Moving forward, we would like to outline several directions for future experimental work aimed at a more refined analysis of the syntax-discourse interface. In order to stay within the scope of our current discussion, we limit these observations to the three accounts that we have chosen to focus on in this article and to the two languages that provided the empirical data for our investigation.[4]

To start, the contribution of syntactic complexity could be further explored by examining the same interface-level properties as discussed above but in different structural contexts. This could be achieved, for example, in an experiment investigating subject and object topics in near-native speakers of Japanese and Korean. These topics belong at the syntax-discourse interface; however, the base position of object topics is less accessible than that of subject topics. If the acquisition of syntactic properties in near-native speakers is mainly unproblematic while the

acquisition of interface properties presents challenges, then we can expect no differences between the subject topic and object topic conditions in highly advanced speakers. If, on the other hand, difficulties with the syntax-discourse interface are a manifestation of greater syntactic complexity, we may expect diminished accuracy on conditions involving object topicalization. The overall pattern of results would then be similar to the findings in relative clauses showing that object gaps are more difficult to process than subject gaps (see Kwon, Kluender, Kutas & Polinsky, 2013 for an overview).

To put another pair of hypotheses to the test, it may be possible to compare Integration Difficulty Hypothesis with the Contextual Embedding Hypothesis. To do so, we would need to create conditions allowing us to examine the same interface phenomenon in contexts that call for varying degrees of cognitive effort, for example, a situation where a given form is associated with several interpretive options versus a context where only one such option is available. A comparison of contrastive topics in root and embedded clauses is a way to test such a distinction. Whereas topics in matrix clauses can either be thematic or contrastive, topics in embedded clauses are almost always contrastive, due to restrictions on the distribution of thematic topics in embedded contexts. Thus, embedded topics are unambiguously contrastive, while the presence of the topic particle in the main clause may signal different interpretations: contrastive or thematic (where thematic can be anaphoric or generic). The ambiguity is of course fully resolvable in context, but resolving it takes some effort, which could lead to diminished performance in bilingual populations under the Contextual Embedding Hypothesis.

## Acknowledgements

This project was supported in part by grants from Center for Advanced Study of Language (CASL) at the University of Maryland and from NSF (SMA-1429961) to the second author. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors. We would like to thank Sun-Hee Bae, Miwako Hisagi, Kenneth Mai, Hiroki Nomoto, Aika

Taguchi, and three anonymous reviewers for helpful comments on this paper. All errors are our responsibility. An earlier, more preliminary version of this paper appeared in MITWPL 13.

## Notes

1. Abbreviations: ACC—accusative; ADN—adnominal; COMP—complementizer; DECL—declarative; GEN—genitive; INF—infinitive; NEG—negation; NOM—nominative; PASS—passive; PRS—present; PST—past; TOP—topic.

2. See also (Laleko & Polinsky, 2013) for some discussion.

3. A full list of materials can be found on the following website: thedata.harvard.edu/dvn/dv/polinsky.

4. We do not *a priori* exclude the possibility of alternative hypotheses, and we look forward to additional cross-linguistic investigations that will expand the data available to researchers today.

## References

**Argyri, E. & Sorace, A.**

(**2007**) Cross-linguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*, 10(7), 79–99. doi: 10.1017/S1366728906002835

**Avrutin, S.**

(**1999**) *Development of the Syntax-Discourse Interface*. Kluwer: Dordrecht. doi: 10.1007/978-94-017-1239-2

**Belletti, A., & Leonini, C.**

(**2004**) Subject inversion in L2 Italian. *EUROSLA yearbook*, 4, 95–118. doi: 10.1075/eurosla.4.06bel

**Benmamoun, E., Montrul, S., & Polinsky, M.**

(**2010**) Prolegomena to heritage linguistics [White paper]. University of Illinois at Urbana-Champaign and Harvard University. Retrieved from www.nhlrc.ucla.edu/pdf/HL-whitepaper.pdf.

(**2013a**) Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, 39, 129–181. doi: 10.1515/tl-2013-0009

(2013b) Defining an "ideal" heritage speaker (a reply to peer commentaries). *Theoretical Linguistics*, 39, 259–294. doi: 10.1515/tl-2013-0018

**Bos, P., Hollebrandse, B. & Sleeman, P.**

(2004) Introduction: The pragmatics-syntax and the semantics-syntax interface in acquisition. *International Review of Applied Linguistics in Language Teaching*, 42, 101–110. doi: 10.1515/iral.2004.004

**Choi, H.W.**

(1999) *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford: CSLI.

**Chomsky, N.**

(1995) *The Minimalist Program*. Cambridge, MA: The MIT Press.

**Chomsky, N. & Miller, G.**

(1963) Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology* (pp. 269–321.) New York: Wiley.

**Donaldson, B.**

(2011) Left dislocation in near-native French. *Studies in Second Language Acquisition*, 33, 399–432. doi: 10.1017/S0272263111000039

(2012) Syntax and discourse in near-native French: Clefts and focus. *Language Learning*, 62(3), 902–930. doi: 10.1111/j.1467-9922.2012.00701.x

**Domínguez, L.**

(2013) *Understanding Interfaces: Second Language Acquisition and First Language Attrition of Spanish Subject Realization and Word Order Variation*. LALD, 55. Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/lald.55

**Fanselow, G. & Frisch, S.**

(2006) Effects of processing difficulty on judgments of acceptability. In G. Fanselow, C. Féry, R. Vogel, & M. Schlesewsky (Eds.), *Gradience in Grammar* (pp. 291–316). Oxford, England: Oxford University Press. doi: 10.1093/acprof:oso/9780199274796.003.0015

**Ferreira, F., & Clifton, C.**

(**1986**) The independence of syntactic processing. *Journal of Memory and Language*, 25, 348–368. doi: 10.1016/0749-596X(86)90006-9

**Frazier, L.**

(**1985**) Syntactic complexity. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives* (pp. 129–189. Cambridge, England: Cambridge University Press. doi: 10.1017/CBO9780511597855.005

**Gibson, E., Piantadosi, S., & Fedorenko, K.**

(**2011**) Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistic Compass*, 5, 509–524. doi: 10.1111/j.1749-818X.2011.00295.x

**Halliday, M.A.K. & Hasan, R.**

(**1976**) *Cohesion in English*. London: Longman.

**Hawkins, R. & Hattori, H.**

(**2006**) Interpretation of English multiple *wh*- questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research*, 22(3), 269–301. doi: 10.1191/0267658306sr269oa

**Heycock, C.**

(**2008**) Japanese -wa, -ga, and information structure. In S. Miyagawa, & M. Saito (Eds.), *The Oxford Handbook of Japanese Linguistics* (pp. 54–83). Oxford University Press.

**Hulk, A. & Müller, N.**

(**2000**) Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism*: *Language and Cognition*, 3(3), 227–244.

**Ivanov, I.**

(**2012**) L2 acquisition of Bulgarian clitic-doubling: A test case for the Interface Hypothesis. *Second Language Research*, 28(3), 345–368. doi: 10.1177/0267658312452066

**Iverson, M., Kempchinsky, P. & Rothman, J.**

(**2008**) Interface vulnerability and knowledge of the subjunctive/indicative distinction with negated epistemic predicates in L2 Spanish. *EUROSLA Yearbook*, 8, 135–163. doi: 10.1075/eurosla.8.09ive

**Iwasaki, S.**

(**1987**) "Identifiability, scope-setting, and the particle wa: A study of Japanese spoken expository discourse." In Hinds, J., Iwasaki, S., & S. Maynard (Eds.), [*Typological studies in language, 14*]. *Perspectives on topicalization: The case of Japanese wa* (pp. 107–142). Amsterdam: John Benjamins. doi: 10.1075/tsl.14.06iwa

**Jackendoff, R.**

(**2002**) *Foundations of Language*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198270126.001.0001

**Kishimoto, H.**

(**2009**) Topic prominency in Japanese. *The Linguistic Review*, 26, 465–513. doi: 10.1515/tlir.2009.017

**Koornneef, A.**

(**2008**) *Eye-Catching Anaphora*. Utrecht: LOT International Dissertation Series.

**Koornneef, A., Avrutin, S., Wijnen, F., & Reuland, E.**

(**2001**) Tracking the preference for bound-variable dependencies in ambiguous ellipses and only-structures. In J. Runner (Ed.), [*Syntax and semantics*, 37]. *Experiments at the interfaces* (pp. 67–100). Bingley, UK: Emerald Group Publishing Limited.

**Kuno, S.**

(**1972**) Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry*, 3(3), 269–320.

(**1973**) *The Structure of the Japanese Language*. Cambridge, MA: The MIT Press.

**Kuroda, S.-Y.**

(**1965**) Causative forms in Japanese. *Foundations of Language*, 1, 30–50.

(**1992**) *Japanese Syntax and Semantics*. Dordrecht: Kluwer. doi: 10.1007/978-94-011-2789-9

(**2005**) Focusing on the matter of topic: A study of wa and ga in Japanese. *Journal of East Asian Linguistics*, 14, 1–58. doi: 10.1007/s10831-004-2701-5

**Kwon, N., Kluender, R., Kutas, M. & Polinsky, M.**

(**2013**) Subject/object processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language*, 89, 537–585. doi: 10.1353/lan.2013.0044

**Laleko, O.**

(**2010**) *The Syntax-Pragmatics Interface in Language Loss: Covert Restructuring of Aspect in Heritage Russian*. Doctoral dissertation, University of Minnesota.

**Laleko, O. & Kawamura, H.**

(**2011**) Grammatical restructuring in intergenerational language shift: *The case of heritage Japanese*. Paper presented at the 56th Annual Conference of the International Linguistics Association. Rutgers, 15–17 April.

**Laleko, O., & Polinsky, M.**

(**2013**) Marking topic or marking case? *Heritage Language Journal*, 10(2), 40–64.

**Langacker, R.**

(**2000**) A dynamic usage-based model. In M. Barlow and S. Kemmer (Eds.), *Usage-Based Models of Language* (pp. 1–64). Stanford: CSLI.

**Leal Méndez, T., Rothman, J. & Slabakova, R.**

(**2015**) Discourse-sensitive clitic-doubled dislocations in heritage Spanish. *Lingua*, 155(4), 85–97. doi: 10.1016/j.lingua.2014.01.002

**Lee, H.S. & Thompson, S.A.**

(**1989**) A discourse account of the Korean accusative marker. *Studies in Language*, 13, 105–128. doi: 10.1075/sl.13.1.04tho

**Lewis, S. & Phillips, C.**

(**2015**) Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*, 44(1), 27–46. doi: 10.1007/s10936-014-9329-z

**Li, C.N. & Thompson, S.A.**

(**1976**) Subject and topic: A new typology of language. In C.N. Li (Ed.), *Subject and Topic* (pp.457–461). Austin: University of Texas Press.

**Lozano, C.**

(**2006**) Focus and split intransitivity: The acquisition of word order alternations in non-native Spanish. *Second Language Research*, 22(2), 145–187. doi:10.1191/0267658306sr264oa

**Machida, N., Miyagawa, S., & Wexler, K.**

(**2004**) A-chain maturation reexamined: why Japanese children perform better on 'full' unaccusatives than on passives. In Csirmaz, A., Gualmini, A., & A. Nevins (Eds.), *MIT Working Papers in Linguistics*, 48 (pp. 91–112). MIT Press, Cambridge, MA.

**Maynard, S.K.**

(**1980**) Discourse Functions of the Japanese Theme Marker 'Wa.' Unpublished doctoral dissertation, Northwestern University.

**Montrul, S.**

(**2002**) Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition*, 5, 39–68. doi:10.1017/S1366728902000135

(**2004**) Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125–142. doi:10.1017/S1366728904001464

(**2005**) Second language acquisition and first language loss in adult early bilinguals: Exploring some differences and similarities. *Second Language Research*, 21(3), 199–249. doi:10.1191/0267658305sr247oa

(**2011**) Multiple interfaces and incomplete acquisition. *Lingua*, 121, 591–604. doi:10.1016/j.lingua.2010.05.006

**Montrul, S., Foote, R., & Perpiñán, S.**

(**2008**) Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning: A Journal of Research in Linguistic Studies*, 58, 503–553. doi:10.1111/j.1467-9922.2008.00449.x

**Müller, N. & Hulk, A.**

(**2001**) Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4(1), 1–21. doi:10.1017/S1366728901000116

**Nakanishi, K.**

(**2001**) Prosody and information structure in Japanese: A case study of topic marker wa. In N. Akatsuka & S. Strauss (Eds.), *Japanese/Korean Linguistics*, 10, 434–447.

**O'Grady, W.**

(**2011**) Interfaces and processing. *Linguistic Approaches to Bilingualism*, 1(1), 63–66. doi: 10.1075/lab.1.1.08ogr

**Ono, H.**

(**2001**) EPP-driven XP movement in Japanese. In Antrim, N.M., Goodall, G., Schulte-Nafeh, M., & V. Samiian (Eds.), *Proceedings of WECOL* 99. California State University: Freshno, CA.

**Pérez-Leroux, A.T. & Glass, W.R.**

(**1997**) "OPC effects on the L2 acquisition of Spanish." In Pérez-Leroux, A.T. & W.R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish*. Vol. I: Developing Grammars (pp. 149–165). Sommerville, MA: Cascadilla Press.

**Phillips, C.**

(**2013**) Parser-grammar relations: We don't understand everything twice. In Sanz, M., Laka, I., & M. Tanenhaus (Eds.), *Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structures* (pp. 294–315.) Oxford, England: Oxford University Press. doi: 10.1093/acprof:oso/9780199677139.003.0017

**Platzack, C.**

(**2001**) The vulnerable C-domain. *Brain and Language*, 77, 364–377. doi: 10.1006/brln.2000.2408

**Polinsky, M.**

(**1997**) American Russian: *Language loss meets language acquisition*.In W. Browne (Ed.), *Annual Workshop on Formal Approaches to Slavic Linguistics: The Cornell Meeting* (pp. 370–406). Ann Arbor: Michigan Slavic Publications.

(**2006**) Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, 14, 191–262.

(**2008a**) Heritage language narratives. In D. Brinton, O. Kagan, & S. Bauckus (Eds.), *Heritage Language Education: A New Field Emerging* (pp.149–164). New York: Routledge.

(**2008b**) Without aspect. In G. Corbett & M. Noonan (Eds.), *Case and Grammatical Relations* (pp. 263–282). Amsterdam: John Benjamins. doi: 10.1075/tsl.81.13pol

(2011) Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition*, 33, 305–28. doi: 10.1017/S027226311000077X

**Polio, C.**

(1995) "Acquiring nothing? The use of zero pronouns by nonnative speakers of Chinese and the implications for the acquisition of nominal reference. *Studies in Second Language Acquisition*, 17, 353–377. doi: 10.1017/S0272263100014248

**Reuland, E.**

(2011) *Anaphora and Language Design*. Cambridge, MA: MIT Press.

**Rizzi, L.**

(1997) The fine structure of the left periphery. In Haegeman, L. (Ed.), *Elements of Grammar. Handbook in Generative Syntax* (pp. 281–337). Dordrecht: Kluwer.

**Rothman, J.**

(2007) Pragmatic solutions for syntactic problems: Understanding some L2 syntactic errors in terms of pragmatic deficits. In Baauw, S., Dirjkoningen, F. & M. Pinto (Eds.), *Romance Languages and Linguistic Theory*(pp. 299–320). Amsterdam: John Benjamins.

(2009) Pragmatic deficits with syntactic consequences?: L2 pronominal subjects and the syntax-pragmatics interface. *Journal of Pragmatics*, 41, 951–973. doi: 10.1016/j.pragma.2008.07.007

**Rothman, J. & Slabakova, R.**

(2011) The mind-context divide: On acquisition at the linguistic interfaces. *Lingua*, 121, 568–576. doi: 10.1016/j.lingua.2011.01.003

**Serratrice, L. Sorace, A. & Paoli, S.**

(2004) Crosslinguistic influence at the syntax-pragmatics interface: Subjects and objects in English-Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition*, 7(2), 183–205. doi: 10.1017/S1366728904001610

**Shibatani, M.**

(1977) Grammatical relations and surface cases. *Language*, 53, 789–809. doi: 10.2307/412912

**Shimojo, M.**

(**2006**) Properties of particle 'omission' revisited. *Toronto Working Papers in Linguistics*, 26, 123–140.

**Shimojo, M. & Choi, H.-W.**

(**2000**) On asymmetry in topic marking: the case of Japanese WA and Korean NUN. In Okrent, A., & J. Boyle (Eds.), *Chicago Linguistic Society*, 36(1) (pp. 455–467). Chicago, IL: Chicago Linguistic Society.

**Slabakova, R. & Ivanov, I.**

(**2011**) A more careful look at the syntax-discourse interface. *Lingua*, 121, 637–651. doi: 10.1016/j.lingua.2010.05.003

**Slabakova, R., Kempchinsky, P. & Rothman, J.**

(**2012**) Clitic-doubled left dislocation and focus fronting in L2 Spanish: A Case of successful acquisition at the syntax-discourse interface. *Second Language Research*, 28(3), 319–343. doi: 10.1177/0267658312447612

**Sorace, A.**

(**2004**) Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition*, 7 (2), 143–145. doi: 10.1017/S1366728904001543

(**2005**) Selective optionality in language development. In Cornips, L. & K.Corrigan, (Eds.), *Syntax and Variation: Reconciling the Biological and the Social* (pp. 55–80). Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/cilt.265.04sor

(**2009**) Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, 13(2), 195–210. doi: 10.1177/1367006909339810

(**2011**) Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism*, 1, 1–33. doi: 10.1075/lab.1.1.01sor

**Sorace, A., & Keller, F.**

(**2005**) Gradience in linguistic data. *Lingua*,115, 1497–1524. doi: 10.1016/j.lingua.2004.07.002

**Sorace, A., & Serratrice, L.**

(**2009**) Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, 13(2), 195–210. doi: 10.1177/1367006909339810

**Sorace, A., Serratrice, L. Filiaci, F., & Baldo, M.**

(2009) Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, 119, 460–477. doi:10.1016/j.lingua.2008.09.008

**Sprouse, J.**

(2011) A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, 43, 155–167. doi:10.3758/s13428-010-0039-7

**Stevenson, S. & Merlo, P.**

(1997) Lexical structure and processing complexity. *Language and Cognitive Processes*, 12(1–2), 349–399. doi:10.1080/016909697386880

**Takezawa, K.**

(1987) A configurational approach to Case-marking in Japanese. Doctoral dissertation, University of Washington.

**Tateishi, K.**

(1994) The syntax of 'subjects.' *Studies in Japanese Linguistics*, 14, 91–108.

**Tomioka, S.**

(2009) Contrastive topics operate on speech acts, In M. Zimmermann & C. Féry (Eds.) *Information Structure: Theoretical, Typological and Experimental Perspectives* (pp. 115–138). Oxford University Press. doi:10.1093/acprof:oso/9780199570959.003.0006

(2010) A scope theory of contrastive topic, *Iberia: International Journal of Theoretical Linguistics*, 2, 113–130.

**Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F.**

(2004) First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English." *International Journal of Bilingualism*, 8(3), 257–277. doi:10.1177/13670069040080030601

**Tsimpli, I. & Sorace, A.**

(2006) Differentiating interfaces: L2 performance in syntax-semantic and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 653–664). Somerville, MA: Cascadilla Press.

**Tsutsui, M.**

(**1984**) *Particle ellipses in Japanese*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

**Ueda, M.**

(**1993**) On the phrase structure of Japanese and English clauses. In N. Hasegawa (ed.), *Japanese Syntax in Comparative Grammar* (pp. 9–44). Tokyo: Kurushio.

**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). Amsterdam: John Benjamins. doi: 10.1075/z.133.16val

**Vermeulen, R.**

(**2005**) Two types of multiple nominative constructions in Japanese. *University of Pennsylvania Working Papers in Linguistics*, 11, 1, 305–319. Retrieved from repository.upenn.edu/pwpl/vol11/iss1/24.

**White, L.**

(**2003**) Fossilization in steady state L2 grammars: Implications of persistent problems with inflectional morphology. *Bilingualism: Language and Cognition*, 6(2), 129–141. doi: 10.1017/S1366728903001081

(**2011**) Second language acquisition at the interfaces. *Lingua*, 121, 577–590. doi: 10.1016/j.lingua.2010.05.005

**Yatabe, S.**

(**1999**) Particle ellipsis and focus projection in Japanese. *Language, Information, Text*, 6, 79–104.

## Appendix

### Appendix: Sample Stimuli

#### I. *Japanese*

(15) Congruous topic/subject marking

a. topic

i. anaphoric

私は先月初めて中国を訪れました。

Watasi-wa sengetu-hazimete tyuugoku-o otozuremas-ita

I-TOP last month-first time China-ACC visit-PAST

'I visited China for the first time last month.'

今、中国は一番好きな国です。

Ima tyuugoku-wa itiban sukina kuni-desu.

now China-TOP first favorite county-is

'Now, China is my favorite country.'

ii. generic

太陽は東から西へ動きます。

Taiyoo-wa higasi-kara nishi-e ugokimasu.

sun-TOP east-from west-to move

'The sun moves from the east to the west.'

iii. contrastive

A) 京子さんはふだん洋楽を聞きますか？

Kyoko-san-wa hudan yoogaku-o kikimasu-ka?

Kyoko-(Ms.)-TOP usual western music-ACC listen-Q?

'Ms. Kyoko, do you usually listen to Western music?'

B) 私はふだん邦楽は聴きますが、洋楽は聴きません。

Watasi-wa hudan hoogaku-wa kikimasu ga yoogaku-wa kiki-masen.

I-TOP usual Japanese music-TOP listen but western music-TOP listen-NEG

'I usually listen to Japanese music but I do not listen to Western music.'

b. subject

i. descriptive

動物園に 2 匹のかわいいライオンの赤ちゃんがいます。
Doobutuen-ni nihiki-no kawaii raion-no akatyan-ga imasu
zoo-in two-GEN cute lion-GEN baby-NOM be
'There are two cute baby lions in the zoo.'

ii. exhaustive

A) 料理に関する本でなにかお勧めありますか？
Ryoori-ni kansuru hon-de nanika osusume arimasu-ka?
cooking-with relate book-at something recommend have-Q?
'Do you have any books about cooking?'

B) この本が一番安くて読みやすいですよ。
Kono hon-ga itiban yasuku-te yomi yasui desu-yo.
this book-NOM best cheap-and read easy be-PTC
'This book is the cheapest and easy to read.'

(16) Incongruous marking

a. topic marker used in place of subject marker
あの車の上を見て！
Ano kuruma-no ue-o mite!
that car-GEN up-ACC look
'Look at the top of the car!'
あそこに小さな猫はいる。
Asoko-ni tiisana neko-wa iru.
there-at small cat-TOP be
'There is a kitty on the car roof.'

b. subject marker used in place of topic marker
昨日の授業参観には多くの保護者たちが出席しました。

Kinoo-no zyugyoosankan niwa ookuno hogosyatati-ga syusseki simas-ita.
yesterday-GEN class observation in-TOP many parents-NOM attend do-PAST
保護者たちが皆、子供と一緒に帰りました。
Hogosyatachi-ga mina kodomo to issyoni kaerimas-ita.
Parents-NOM all children with together go back-PAST
'Many parents attended yesterday's class observation. All of the parents went home with their children.'

(17) Omission of topic marker

a. congruous
人口一番多い都市東京
Jinkoo ichiban ooi tosi Tookyo
Population first many city Tokyo
'Tokyo (is) the city which has the biggest population (in Japan).'

b. incongruous
光、音より伝わるスピードが早いです。
Hikari oto yori tutawaru spiido hayai desu.
Light sound than transmit speed fast be
'Light is faster than sound.'

(18) Omission of subject marker

a. congruous
最近始まった人気ドラマ映画になるんだって。
Saikin hazima-tta ninki dorama eiga ni naru-n da-tte.
Recently start-PAST popular drama movie to become-NMLZ be-seem
'Apparently the new popular drama which started recently will become a movie.'

b. incongruous
大会で有名な陸上選手走った

Taikai-de yumeina rikuzyoo-sensyu hasi-tta.
Competition-at famous runner run-past
'A famous runner ran in the competition.'

## II. *Korean*

(15) Congruous topic/subject marking:

a. topic

i. anaphoric

나는 저번 달에 처음으로 중국에 갔어.

Na-nun cepe-n tal-ey cheum-ulo cwungkwuk-ey ka-ss-e.
I-TOP last-ADN month-in first_time-ADV China-LOC go-PAST-DECL
'Last month, I went to China for the first time.'

이제 내가 제일 좋아하는 나라는 중국이야.

icey nay-ka ceyil cohaha-nun nala-nun cwungkwuk-iya.
Now I-NOM most like-ADN country-TOP China-DECL
'Now the country I like the most is China.'

ii. generic

태양은 동쪽에서 서쪽으로 움직입니다.

Taeyang-eun dong-jjock-aesuh suhjjock euro umjickyipnida.
Sun-TOP east-part-from west-part to move-DECL
'The sun moves from the east to the west.'

iii. contrastive

A) 당신은 국악을 좋아하나요?

tangsin-un kwukak-ul cohaha-nayo?
you-TOP traditional music-ACC like-Q?
'Do you like traditional music?'

B) 팝송은 좋아하지만 국악은 별로 좋아하지 않아요.
phapsong-un cohaha-ciman kwukak-un pyello cohahaci anh-ayo.
pop song-TOP like-but traditional music-TOP much like NEG-DECL
'I like pop songs, but I don't like traditional music very much.'

b. subject

i. descriptive
동물원에는 귀여운 아기 사자 두 마리가 있다.
Tongmwulwen-ey-nun kwiyew-un aki saca twu mali-ka iss-ta.
Zoo-LOC-TOP cute-ADN baby lion two CL-NOM be-DECL
'There are two cute baby lions in the zoo.'

ii. exhaustive
A) 한국 역사-에 관한 책 추천해줄 수 있어?
hankwuk yeksa-ey kwanha-n chayk chwuchenhay cwul swu is-se?
Korea history-to about-ADN book recommend give DEL exist-Q?
'Can you recommend any books about Korean history?'

B) 이 책-이 제일 싸고 내용도 별로 안 어려워.
i chayk-i ceyil ssa-ko nayyong-to pyello an elye-we.
this book-NOM most cheap-and content-also particularly not hard-DECL
'This book is the cheapest and not particularly hard to read.'

(16) Incongruous marking

a. topic marker used in place of subject marker
저 자동차 위를 봐!
ce catongcha wi-lul p-wa!
'Look at the top of the car!'
저기에 작은 고양이는 있어.

ceki-ey cak-un koyangi-nun iss-e.
there-LOC small-ADN cat-TOP exist-DECL
'There is a small cat there.'

b. subject marker used in place of topic marker
어제 말했던 재미있는 만화책 말인데,
ecey malha-yss-ten caymiiss-nun manhwachayk malintey,
yesterday talk-PAST-ADN interesting-ADN comic book talk
그 책-이 어디에서 팔아?
ku chayk-i eti-eyse phal-a?
that book-NOM where-LOC sell-Q?
'Speaking of the interesting comic book that we talked about yesterday, where is that book sold?'

(17) Omission of topic marker

a. congruous
어제 밤에 너랑 본 영화, 정말 재미있었어.
ecey pam-ey ne-lang po-n yenghwa, cengmal caymiiss-ess-e.
yesterday night-in you-with see-ADN movie very interesting-PAST-DECL
'The movie which I watched with you yesterday night was very good.'

b. incongruous
빛 소리-보다 전파 속도가 빠르다.
pich soli-pota cenpha s okto-ka ppalu-ta
light sound-than transmission speed-NOM fast-DECL
'Light is faster than sound in terms of transmission speed.'

(18) Omission of subject marker

a. congruous
최근 유행인 그 드라마,

choykun yuhayng-in ku tulama,
recently popular-ADN that drama,
결국 영화로도 제작된다고 들었어.
kyelkwuk yenghwa-lo-to ceycaktoyn-ta-ko tul-ess-e.
ultimately movie-as-also make-DECL-that hear-PAST-DECL
'That drama that has been popular recently, I heard that it will also be made into a movie.'

b. incongruous
우리 가족들 중에서 유일하게 아버지
wuli kacok-tul cwung-eyse yuilha-key apeci
our family-PL in-from only-ADV father
내가 유학가는 것을 허락하셨다
nay-ka yuhakka-nun kesul helakha-sy-ess-ta.
I-NOM study abroad-ADN thing-ACC permit-HON-PAST-DECL
'Among our family members, only my father permitted me to study abroad.'

## Authors' addresses

**Oksana Laleko** **Maria Polinsky**

Linguistics Program

JFT 714, 600 Hawk Drive

SUNY New Paltz

New Paltz, NY 12561 USA

lalekoo@newpaltz.edu