RESEARCH

Cross-linguistic scope ambiguity: When two systems meet

Gregory Scontras[1], Maria Polinsky[2], C.-Y. Edwin Tsai[3] and Kenneth Mai[4]

[1] University of California, Irvine, US

[2] University of Maryland, US

[3] City University of Hong Kong, HK

[4] Harvard University, US

Corresponding author: Gregory Scontras (g.scontras@uci.edu)

Accurately recognizing and resolving ambiguity is a hallmark of linguistic ability. English is a language with scope ambiguities in doubly-quantified sentences like *A shark ate every pirate*; this sentence can either describe a scenario with a single shark eating all of the pirates, or a scenario with many sharks—a potentially-different one eating each pirate. In Mandarin Chinese, the corresponding sentence is unambiguous, as it can only describe the single-shark scenario. We present experimental evidence to this effect, comparing native speakers of English with native speakers of Mandarin in their interpretations of doubly-quantified sentences. Having demonstrated the difference between these two languages in their ability for inverse scope interpretations, we then probe the robustness of the grammar of scope by extending our experiments to English-dominant adult heritage speakers of Mandarin. Like native speakers of Mandarin, heritage Mandarin speakers lack inverse scope in Mandarin. Crucially, these speakers also lack inverse scope in English, their dominant language in adulthood. We interpret these results as evidence for the pressure to simplify the grammar of scope, decreasing ambiguity when possible. In other words, when two systems meet—as in the case of heritage speakers—the simpler system prevails.

Keywords: scope ambiguity; English; Mandarin Chinese; heritage speakers

1 Introduction

Quantifier scope ambiguities have stood at the heart of linguistic inquiry for decades. Montague (1973) builds the possibility for scope-shifting into his seminal work in semantics. May (1977) proposes the rule of QR, which derives scope ambiguities syntactically. Both the semantic and syntactic approaches ensure that doubly-quantified sentences are ambiguous, as in (1). Viewing quantifiers like *a* and *every* as logical operators, the ambiguities correspond to the relative scope of these operators within the logical form (LF) of the sentence (whence the name "scope ambiguities").

(1)   A shark attacked every pirate.
      a.   SURFACE SCOPE ($\exists > \forall$):
          There was a single shark that attacked multiple pirates.
      b.   INVERSE SCOPE ($\forall > \exists$):
          For each pirate, there was a (different) shark that attacked him.

Figure 1 provides verifying scenarios for the two readings of the sentence in (1).[1] On the left, we have a single shark attacking all of the pirates, a single event corresponding to the SURFACE interpretation of the sentence. On the right, we have multiple shark attacks, one for every pirate, corresponding to the INVERSE interpretation of the sentence. The SURFACE interpretation of (1) is true only in the left panel of Figure 1; in the right panel of Figure 1 there is no single shark that attacked every pirate. Note, however, that the INVERSE interpretation is true in both panels of Figure 1: on the right, every pirate has a shark attacking him. Similarly in the left panel: every pirate has a shark attacking him; the shark just happens to be the same. We return to the logical relationship between scope interpretations presently, in Section 2.

For speakers of English, such intuitions about scope ambiguity for doubly-quantified sentences are stable and readily accessible. However, not every language is like English, and we have no reason to suspect that every language affords the same range of interpretations to its doubly-quantified sentences. Our first task is to demonstrate that languages vary with respect to scope ambiguity, comparing English with Mandarin Chinese, a language widely believed to lack scope ambiguity (Huang 1981; 1982; Lee 1986; Aoun & Li 1989; 2003), but one that has not been investigated using the same experimental paradigms as English. To better understand the prohibition on inverse scope in Mandarin, and, conversely, its possibility in English, we then shift our sights to cases where the two systems meet: English-dominant heritage speakers of Mandarin, as well as heritage speakers of English. But first, we will briefly examine the native grammars, building on the large body of existing research, as well as our own earlier work.

2 Background

Ours is not the first investigation of quantifier scope ambiguities. To begin, we review the relevant theoretical and experimental literature on native English judgments. We then turn to native speakers of Mandarin, highlighting a recent controversy over the status of scope ambiguities in this language—a controversy we aim to resolve with the results of our experiments. Finally, we introduce heritage language study—the investigation of early simultaneous and/or sequential bilinguals dominant in a language other than their first/home language—and discuss its relevance to the topic at hand.

Figure 1: Scenarios verifying the surface (left) and inverse (right) interpretations of the sentence in (1).

---

[1] All images, which were also used in our experiment materials, come from Benjamin Bruening's Scope Fieldwork Project (http://udel.edu/~bruening/scopeproject/scopeproject.html). Although verifying scope interpretations with pictures may have its limitations, this method has proven to be an effective means of establishing scope readings and differences across languages (e.g., Bruening 2008; Bochnak & Matthewson 2015).

2.1 English

English sentences with more than one quantificational expression exhibit scope ambiguities. The ambiguities correspond to the relative scoping of the quantificational expressions at logical form. For present purposes, we limit our discussion to doubly-quantified sentences as in (1), repeated in (2), and (3).

(2)   A shark attacked every pirate.
   a.   SURFACE SCOPE ($\exists > \forall$):
        There was a single shark that attacked each pirate.

   b.   INVERSE SCOPE ($\forall > \exists$):
        For each pirate, there was a (different) shark that attacked him.

(3)   Every shark attacked a pirate.
   a.   SURFACE SCOPE ($\forall > \exists$):
        For each shark, there was a (different) pirate that it attacked.

   b.   INVERSE SCOPE ($\exists > \forall$):
        There was a single pirate that was attacked by each shark.

Before we consider proposals for the generation of the ambiguities, it bears emphasizing that the SURFACE and INVERSE interpretations of doubly-quantified sentences are not logically independent of each other (as pointed out by, e.g., Reinhart 1976; 1997; Cooper 1979; Ruys 1992). Specifically, the INVERSE interpretation of (3)—a sentence with universal *every* scoping over existential *a* at surface structure—entails the SURFACE interpretation: if there was a single pirate that every shark attacked, then necessarily every shark attacked a pirate. For this reason, sentences with *every* in subject position and *a* in object position are poor test cases for the availability of inverse scope. Whenever we say of such sentences that they have received an INVERSE interpretation, in fact the SURFACE interpretation will also hold true via entailment.[2]

In (2)—a sentence with existential *a* preceding universal *every* at surface structure—the SURFACE interpretation entails the INVERSE: if there was a single shark that attacked each pirate, then necessarily each pirate was attacked by a shark. Because the INVERSE interpretation of (2) does not entail its surface interpretation, here we have the test case for inverse scope: for people to accept (2) as a description of the INVERSE-satisfying scenario (i.e., the right panel of Figure 1 above), they must have given the sentence an inverse interpretation. Thus, for the purpose of evaluating inverse scope in doubly-quantified sentences with universal *every* and existential *a*, responses to sentences like (1), where *a* precedes *every*, will be of primary interest.

While speakers of English often accept inverse interpretations of doubly-quantified sentences, they also display a reliable and robust preference for surface interpretations. This preference holds across a variety of dependent measures (e.g., measures of grammaticality like sentence ratings and truth judgments, or measures of processing difficulty like reaction/reading times), at a range of ages. Various proposals have been put forth to explain this preference, and they all share the feature that INVERSE scope calculation is costly relative to *surface* scope. This cost delivers the observed preference for SURFACE scope. Anderson (2004) identifies this cost as relating to the syntax of scope configurations, proposing the Processing Scope Economy principle.

---

[2] Because of this entailment relation between INVERSE and SURFACE interpretations of sentences such as (3), Mayr & Spector (2011) appeal to economy to rule out INVERSE scope in the first place. Under such an approach, (3) is unambiguous—another reason why these sentences are a poor test case for inverse scope.

(4)   *Processing Scope Economy* (Anderson 2004: 48):
        The human sentence processing mechanism prefers to compute a scope
        configuration with the simplest syntactic representation (or derivation). Com-
        puting a more complex configuration is possible but incurs a processing cost.

Assuming that scope interpretations correspond to distinct LFs, the principle in (4) privi-
leges surface scope by assuming that the derivation of the INVERSE LF requires more
effort. That is, INVERSE scope involves a more complex syntactic derivation.

   Anderson ties this complexity to the additional syntactic computation that she
assumes generates inverse scope interpretations: QR (May 1977; 1985). Under a QR
approach, the SURFACE and INVERSE interpretations of (2) follow from the schematic
LFs in (5-a) and (5-b), respectively. (Here we simplify Anderson's derivations, ignor-
ing the initial movement of the subject that follows from the VP-internal-subject
hypothesis.)

(5)   A shark attacked every pirate.
        a.   SURFACE SCOPE ($\exists > \forall$):
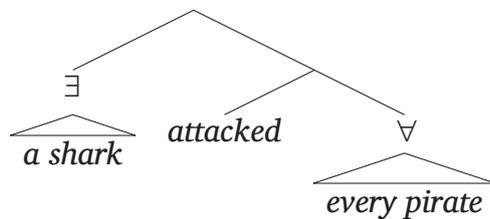
        b.   INVERSE SCOPE ($\forall > \exists$):
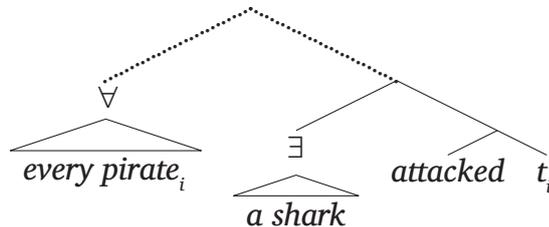
In (5-a), the SURFACE interpretation follows from a basic LF, which preserves the scope
relations from surface structure (i.e., $\exists > \forall$). The INVERSE LF in (5-b) involves an addi-
tional step, covert QR of the object *every pirate* above the subject *a shark*. It is this addi-
tional operation of QR that penalizes the INVERSE LF, and thus the INVERSE interpreta-
tion.

   While we are not committed to a specific approach deriving scope ambiguity, we adopt
this analysis for consistency as we consider the varying scope possibilities and the mecha-
nisms that deliver them across grammars.[3] It bears noting that Anderson's principle of
Processing Scope Economy likely interacts with other pressures to simplify scope calcula-
tions. For example, Kurtzman & MacDonald (1993) follow Fodor (1982) in proposing a
single reference principle: listeners build an on-line parse of the sentences they hear; when
they encounter a singular indefinite at the start of a sentence, they imagine and commit to
just a single referent associated with it (e.g., a single shark). This single-referent parse is
at odds with a many-referent scenario, as in the right panel of Figure 1 above, providing

---

[3] A popular, non-transformational alternative to QR is based on choice functions (Reinhart 1997; Winter
   1997; Kratzer 1998). Recent work by Bergen & Goodman (in prep.) uses implicit quantificational domain
   restriction to derive scope shifting via pragmatic inference. As with QR, these approaches all attribute a
   greater cost to inverse interpretations.

yet another reason why speakers would prefer a surface interpretation. We return to the pressures to simplify scope calculations in the discussion of our experimental findings below. For now the takeaway is that English appears to allow inverse interpretations, in spite of pressures to avoid them.

2.2 Mandarin

In contrast to the permissiveness of English scope calculations, the picture in Mandarin appears remarkably stark. Since the seminal work of Huang (1982), many linguists have arrived at or accepted the conclusion that Mandarin does not allow inverse scope in doubly-quantified sentences (see also Huang 1981; Lee 1986; Aoun & Li 1989; 2003). This prohibition means that Mandarin translations of the English sentences we have so far considered reportedly allow only a surface interpretation. With respect to the scenarios depicted in Figure 1, (6-b) should therefore be judged true only with respect to the left-hand image, and false with respect to the right-hand one.

(6)   a.   Mei-yi-tiao   shayu dou gongji-le   yi-ge   haidao.
              every-one-CLF   shark   all   attack-PST   one-CLF   pirate
              'Every shark attacked a/one pirate.'   ($\forall > \exists$ *only*)

        b.   You   yi-tiao   shayu gongji-le   mei-yi-ge   haidao.
              exist one-CLF   shark   attack-PST every-one-CLF   pirate
              'A/one shark attacked every pirate.'   ($\exists > \forall$ *only*)

To account for the lack of inverse scope interpretations in doubly-quantified sentences in Mandarin, Huang (1982) proposes what Aoun & Li (1989) term the "Isomorphic Principle," which rules out inverse scope by disallowing the LFs that would deliver it.

(7)   *The Isomorphic Principle* (Huang 1982; Aoun & Li 1989):
        Suppose A and B are Quantifier Phrases. Then if A c-commands B at
        S(urface)-Structure, A c-commands B at LF.

As the name suggests, the principle in (7) mandates a strict isomorphism between scope relations at surface structure and scope relations at LF. Note that the principle does not rule out syntactic operations like QR, but, crucially, it does ensure that QR is scope-preserving. Specifically, the LF in (5-b) above violates this principle because *a shark* c-commands *every pirate* before QR at S-structure, but not after QR at LF.

   While the facts the Isomorphic Principle is meant to characterize—namely, the lack of inverse scope interpretations for doubly-quantified sentences in Mandarin—have been more or less unchallenged since they were originally put forth (but see Aoun & Li 1989 for relevant discussion of the scope possibilities for passive sentences), a recent experimental study of Mandarin scope by Zhou & Gao (2009) called them into question. To our knowledge, theirs is the first experiment systematically investigating judgments about scope in Mandarin, and their conclusion comes as a drastic departure from the received wisdom on Mandarin scope: according to Zhou & Gao, doubly-quantified sentences in Mandarin *do* allow inverse scope.

   To evaluate the possibility for inverse scope in Mandarin, Zhou & Gao ran an acceptability rating study with native speakers of Mandarin in Beijing. Participants were provided with one of two possible context scenarios, followed by a doubly-quantified test sentence. Participants were instructed to rate on a five-point scale how well the test sentence described the context scenario that preceded it. The scenarios were meant to verify either a surface or an inverse interpretation of the test sentence. Zhou & Gao's results demonstrate

that, consistent with the observed preference for surface scope (in English) discussed above, participants rated surface conditions significantly higher than inverse conditions. However, inverse conditions received an average rating above 3 (out of 5), a rating which the authors take as evidence for the availability of inverse scope in these sentences.

There stands a major obstacle to Zhou & Gao's conclusion, and it concerns the type of doubly-quantified sentences for which they elicited judgments. All of their test sentences were of the form in (8), where universal *mei* 'every' scopes over existential *yi* 'one' at S-structure.

(8)   Mei-ge   ren   dou qu-le   yi-jia   gongchang.
        every-CLF person all   go-PST one-CLF factory
        'Everyone went to a factory.'

In the inverse condition, subjects rated (8) as a description of the scenario in (9).

(9)   INVERSE *context scenario*:
        "Last summer vacation XiaoZhang, XiaoLi and XiaoWang didn't go home.
        They all took a part time job in the clothing factory near the University."

For (8) to be judged true in the scenario in (9), the authors reason, the sentence must receive an inverse interpretation: there was a single factory such that every person went to it. But, as we and many of our predecessors have stressed, a surface interpretation of (8) also holds true in (9): for each person, there is a factory (that the person went to); the factories happen to be the same. Again, judgments on so-called inverse interpretations of sentences where a universal precedes an existential quantifier at S-structure cannot confirm the possibility of inverse scope, as the inverse interpretation entails the surface. This fact calls into question the findings from Zhou & Gao (2009), and leaves unsettled the scope behavior of doubly-quantified sentences in Mandarin.

While the theoretical literature stemming from Huang (1982) has arrived at the general consensus that Mandarin lacks inverse scope, the experimental data are lacking. Given the problems with the study by Zhou & Gao (2009), our first task is to test the possibility of inverse scope in Mandarin. Our approach will be to compare English, a language with demonstrated inverse scope for doubly-quantified sentences, with Mandarin, a language whose scope calculus is in question. For this comparison to succeed, we must ensure that our participants are responding to tests of like things. In other words, we must be sure that our materials represent faithful translations from one language to the other.[4] Even without appeal to Quine's indeterminacy of translation—the troubles associated with translating complex ideas or concepts across languages—this task proves a difficult one.

In (10) and (11) we present the representative English sentences side-by-side with their Mandarin counterparts. For ease of reference, we adopt the label "$\forall > \exists$" for sentences as in (10) where a universal precedes an existential at S-structure, and "$\exists > \forall$" for sentences as in (11) where an existential precedes a universal.

(10) $\forall > \exists$:
        a.   Every shark attacked a pirate.
        b.   Mei-yi-tiao   shayu dou gongji-le   yi-ge   haidao.
              every-one-CLF shark all   attack-PST one-CLF pirate
              'Every shark attacked a/one pirate.'

---

[4] Mandarin is associated with one of the famous cases of controversy raised by inadequate translation; the initial findings suggesting that Mandarin speakers have trouble with counterfactual interpretations were based on inadequate translations of English sentences (Au 1983; Yeh & Gentner 2005).

(11)   ∃ > ∀:
    a. A shark attacked every pirate.
    b. You   yi-tiao   shayu   gongji-le   mei-yi-ge   haidao.
        exist one-CLF shark attack-PST every-one-CLF pirate
        'A/one shark attacked every pirate.'

Here it bears noting two properties of the Mandarin sentences: first, the Mandarin indefinite expression *yi* 'a' serves double duty as the numeral 'one', and second, sentence-initial indefinite phrases like *yi-tiao-shayu* 'one/a shark' require the existential predicate *you* introducing them. The word *you*, literally 'have', is used to form the existential construction, with the basic structure [*you* DP XP] (Huang 1987; Liu 2011). *You* appears in the initial position of the sentence, as shown in (11-b). The post-*you* NP (*shayu* 'shark' in (11-b)) is usually followed by a predicate-like phrase XP, as in our example (11). *You* can also appear as a predicate in locative/temporal and possessive clauses. The resulting structure of the Mandarin ∃ > ∀ configuration might then be bi-clausal, with *you* composing with an indefinite object which is then modified by a relative clause containing a universally quantified noun, as schematized in (12).

(12)   *Possible bi-clausal structure for existential* you:
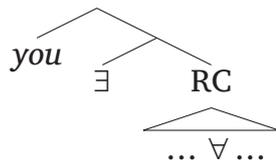
While the syntax and semantics of English indefinites might yet prove elusive, this syntax does not look anything like the structure in (12). We therefore ought to consider a potentially more plausible structure to match the Mandarin ∃ > ∀ sentences, namely existential *there* constructions as in (13).

(13)   There is a shark that attacked every pirate.

Returning to indefinite *yi* 'a/one', here it bears noting Mandarin does not have an article system. We translate Mandarin *yi* as English *a*, but it is not obvious whether the numeral *yi* 'one' is genuinely ambiguous between an indefinite article and a true numeral. We therefore do not know whether *yi* contributes merely existential force (like *a*), or whether it behaves always as a full-fledged numeral (like *one*). Owing to this uncertainty, we ought also to consider English sentences where numeral *one* serves instead of the article *a*. Sticking to sharks and pirates, in (14) and (15) we present the resulting possible English translations for the Mandarin sentences.

(14)   ∀ > ∃:
    a. Every shark attacked a pirate.
        Every shark attacked one pirate.
    b. Mei-yi-tiao   shayu dou gongji-le yi-ge   haidao.
        every-one-CLF shark all   attack-PST one-CLF pirate
        'Every shark attacked a/one pirate.'

(15)   ∃ > ∀:
    a. A shark attacked every pirate.
        One shark attacked every pirate.
        There is a shark that attacked every pirate.
        There is one shark that attacked every pirate.

  b. You yi-tiao shayu gongji-le mei-yi-ge   haidao.
      exist one-CLF shark attack-PST every-one-CLF pirate
      'A/one shark attacked every pirate.'

In Expts. 1 and 2, we use materials like those given above to compare the possibility for inverse scope in English and Mandarin. To be clear: considering these variants of the English sentences is not a commitment on our part to any particular analysis of the Mandarin sentences, but rather due diligence in our pursuit of an apt comparison of doubly-quantified sentences in these two languages.

  Settling the controversy surrounding inverse scope in Mandarin allows us to set the stage for the ultimate aim of this study. Finding (as we do) a difference between native English and Mandarin grammars as they relate to inverse scope, our focus then shifts to the source and stability of this difference. To evaluate these issues, we consider yet another grammar: that of heritage speakers.

2.3 Heritage language speakers

Since its inception, the generative tradition within linguistic theory has concerned itself primarily with monolingual speakers in its quest for what we know when we know (a) language. Chomsky provides an early characterization of the enterprise, focusing attention on idealized language users:

> "Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance."
>
> (Chomsky 1965: 3)

The rapid ascension of formal linguistics over the intervening five decades has demonstrated the success of this approach to the study of language. However, this approach necessarily excludes a wide swath of the world's language users, communities, and even languages. With progress in linguistics, we are now in a much better position to include multilingual speakers in the empirical base of linguistics more generally, and theoretical linguistics in particular. Here we focus on a subset of multilingual language users: heritage speakers. These are simultaneous or sequential unbalanced bilinguals, whose home (minority) language is the weaker of the two (cf. Rothman 2009; Benmamoun et al. 2013a; b; Kupisch 2013; Scontras et al. 2015; Kupisch & Rothman 2016; Montrul 2016). Heritage languages, whose speakers are numerous and widely available, present a unique testbed for issues of acquisition, maintenance/robustness, and transfer within linguistic theory.

  In their expansive overview, Benmamoun et al. (2013b) provide a working profile of heritage speakers. Heritage speakers grow up hearing and speaking both the heritage language (L1) and the majority language (L2). At or around the onset of schooling in the majority language, the majority language becomes the heritage speaker's primary language, supplanting the heritage language as the speaker's dominant language. As a result, proficiency in the heritage language weakens. The study of heritage languages thus stands to identify those areas of grammar that are susceptible to attrition, and those that are not (Benmamoun et al. 2013b; Lohndal 2013).

  The weakening of heritage language evidences similar patterns of decay across a variety of speech communities. According to Benmamoun et al. (2013b: 153), "phonology, in general, seems to be the best-preserved area of the heritage grammar, followed by syntax, while inflectional morphology, semantics, and the syntax-discourse interface are the

most vulnerable." Most susceptible, then, are those areas of grammar that implicate an interface between linguistic modules or levels of representation (Polinsky 2011; Sorace 2011; Pascualy Cabo et al. 2012); and it is precisely at this interface that quantifier scope ambiguities reside.

Scope interpretations bring together at least three levels of representation: syntax (expressing the structural relationship among quantifiers), semantics (expressing the logical implications of this structure), and pragmatics (resolving the ambiguity in context). We might therefore expect scope calculations to diverge from the native grammar in heritage speakers as they perform the costly operation of integrating these various levels of linguistic representation. This divergence could take one of two paths: transfer from the dominant language resulting in an otherwise uncharacteristic pattern of behavior in the heritage speaker; or, faced with two systems of relatively different complexity, the simpler system winning out in the heritage grammar. (Alternatively, there could be no divergence between heritage and native speakers, demonstrating the robustness of the grammar of scope.)

The grammar of scope ambiguities has not received extensive attention in heritage language studies. Lee et al. (2011) test English-dominant heritage speakers of Korean on the interpretation of negative sentences with universally quantified objects, as in (16). In English, this configuration yields ambiguity, corresponding to the scope of negation with respect to the universal quantifier; we use the SURFACE vs. INVERSE terminology to describe the ambiguity.

(16)   Mary didn't read all the books.
   a.   SURFACE SCOPE ($\neg > \forall$):
        It is not the case that Mary read all the books.

   b.   INVERSE SCOPE ($\forall > \neg$):
        For each book, it is not the case that Mary read it.

Despite the availability of both surface and inverse interpretations for sentences like (16), speakers of English demonstrate a strong preference for surface interpretations. Presented with contexts supporting one or the other interpretation, native speakers of English accept inverse interpretations approximately 50% of the time (compared with a 90% acceptance rate for surface interpretations; Lee 2009).[5]

In Korean, similar sentences yield the opposite preference for interpretations (Han et al. 2007; O'Grady et al. 2009). Testing native speakers on sentences as in (17), Lee et al. (2011) show that ostensibly surface interpretations yield near–50% acceptance rates, while inverse interpretations are accepted 90% of the time—the reverse of the English pattern.

(17)   *Korean*:
        Mary-ka   motun chayk-ul   am   ilk-ess-ta.
        Mary-NOM   all   book-ACC   not   read-PST-DECL
        'Mary didn't read all the books.'

Citing a processing explanation of these preferences from Grodner & Gibson (2005), Lee et al. suggest that differences in word order between English and Korean deliver the diverging patterns. In English, an incremental parser first encounters the negative auxiliary *didn't*, followed by the universally quantified object. As the parser encounters each

---

[5] Bear in mind that negative sentences with universal quantifiers present problems with logical entailment similar to our doubly-quantified sentences: the inverse interpretation of (16) entails the surface interpretation. However, given the interest in inverse interpretations, this entailment relation does not alter our or the authors' conclusions.

element, it immediately assigns an interpretation, resulting in the $\neg > \forall$ parse. Generating an inverse interpretation requires revising the initial parse, disrupting the linear operation of the parse and incurring a cost that results in a preference against the inverse, non-linear $\forall > \neg$ interpretation Moreover, this inverse interpretation follows unambiguously from a readily-available alternative utterance: *Mary didn't read any books* (cf. the "pragmatic calculus" of Musolino & Lidz 2006). In Korean, the SOV word order has this same parser first encounter the universally quantified object, then negation; using the same reasoning used for English, here we predict the opposite preference, namely a preference for inverse interpretations in Korean.

The question then becomes: what happens when these systems of preferences meet? Lee et al. used similar materials—negative sentences with universally quantified objects—to test the interpretation preferences of English-dominant heritage speakers of Korean *in English*. Their results show that these heritage speakers deploy their Korean preferences in English: 50% acceptance rate for surface vs. 90% for inverse. Perhaps surprisingly, early exposure to Korean seemed to interfere with scope calculation in English.

Whatever its explanation, this result nevertheless raises important questions concerning the representation of scope in both monolingual and bilingual speakers. What aspect of the dominant English grammar was affected by Korean? Unfortunately, Lee et al. did not test the scope preference of their heritage subjects in the Korean grammar. Since that language was, at the time of the study, the weaker of the two in the subjects' bilingual representation, it is important to determine whether the scope preferences observed in monolingual Korean are still present in that language when it is weakened by a dominant L2. The present study addresses these concerns by testing English-dominant heritage speakers of Mandarin in both English and Mandarin. But there is another, more important difference between our study and that of Lee et al. (2011).

Lee et al. demonstrate diverging preferences of scope interpretations between Korean and English in negative sentences with universally-quantified objects. Crucially, speakers of each language allow both surface and inverse interpretations of these sentences, thus they merely prefer one interpretation over the other. As in the case of doubly-quantified sentences in English, this preference manifests as ~50% acceptance rate for the dispreferred interpretation. However, if Mandarin truly disallows inverse scope in doubly-quantified sentences, here we face a fundamentally different comparison: one language whose grammar permits inverse scope (English) versus another whose grammar does not (Mandarin). This comparison allows us to more directly probe the robustness of each system as they intersect in the heritage grammar.

2.4 Summary and outlook

Research on quantifier scope ambiguities in English demonstrates the viability of both surface an inverse interpretations for doubly-quantified sentences. A similarly clear picture results for Mandarin, whose doubly-quantified sentences are claimed to disallow inverse interpretations. The current study considers what happens when these two grammars intersect in heritage speakers. To that end, we begin our investigation by validating our experimental paradigm and establishing the facts in the native grammars. We then shift focus to heritage speakers, using the same materials to assess the robustness of the grammar of scope, that is, whether scope calculations are susceptible to attrition or transfer from another language.

3 Testing the native grammars

We used an acceptability-rating task to investigate the scope interpretations of English and Mandarin sentences with an existential and a universal quantifier. We follow recent recommendations in the study of heritage languages and employ a gradient acceptability-

rating task, as opposed to a binary truth-judgment task, at the latter paradigm proves unnecessarily taxing for heritage speakers (Laleko & Polinsky 2013; 2016; Montrul 2016; Orfitelli & Polinsky 2017). Given that our ultimate aim is heritage speakers who lack schooling and therefore literacy in their first language (Benmamoun et al. 2013b), the test sentences are presented orally with verifying scenarios that are visual scenes.

3.1 Experiment 1: Native English

We begin with a look at adult native speakers of English tested on their interpretations of English doubly-quantified sentences. We split this experiment into four sub-experiments, according to whether the head of the singular indefinite nominal was the article *a* or the numeral *one*, and whether sentences in the $\exists > \forall$ configuration participated in a *there*-existential.

3.1.1 Participants

We recruited 130 participants via Amazon.com's Mechanical Turk crowd-sourcing service. 114 participants indicated that they were native speakers of English; only their data were included in the analyses reported below. All participants were compensated for their participation.

3.1.2 Design, methods, and materials

Participants took the experiment online using the web-based experiment platform ExperigenRT (Becker & Levine 2010; Pillot et al. 2012). They began by filling out a demographic questionnaire, then completed a training session consisting of three slides. The training items served to ensure that the audio played (and was heard) and that pictures were correctly displayed. Training items also ensured that participants understood the instructions and the correspondence between the sentence and the picture.

In each trial, a picture appeared on the screen and participants clicked on an audio button below the image to play a recorded sentence. After hearing the sentence, participants were asked to judge whether the sentence they heard appropriately described the picture using a 7-point Likert scale (1 = "completely inappropriate", 7 = "completely appropriate"). Participants completed 20 trials in a random order (7 critical items and 13 fillers). Only one version of each test item was presented to any given participant; conditions were chosen at random, with the constraint that participants encountered a given condition from a single sub-experiment at most one time.

Stimuli consisted of audio sentence-picture pairs. Test items featured doubly-quantified sentences with a universal quantifier (*every*) and an existential quantifier (*a, one*). Sentences were recorded by an adult male native speaker of American English with neutral intonation. Pictures co-occurring with sentences verified either a surface or an inverse interpretation of the sentence (with the caveat concerning entailment relations discussed above).

We manipulated two factors: surface ORDER ("$\exists > \forall$" vs. "$\exists > \forall$") and INTERPRETATION ("surface" vs. "inverse") of the sentence. In an attempt to match the language-specific properties of Mandarin in our English sentences, we used four sets of English constructions as targets for translation.[6]

(18)   *Sample sentences from each sub-experiment*:
    a.   PLAIN
        Every shark attacked a pirate.   ($\forall > \exists$)
        A shark attacked every pirate.   ($\exists > \forall$)

---

[6] See the Appendix for the full set of English test sentences.

b.   ONE
     Every shark attacked one pirate.   (∀ > ∃)
     One shark attacked every pirate.   (∃ > ∀)

c.   THERE
     There is a shark that attacked every pirate.   (∃ > ∀)

d.   THEREONE
     There is one shark that attacked every pirate.   (∃ > ∀)

In the PLAIN sub-experiment, (18-a), sentences featured the article *a* and a simple transitive frame. In the ONE sub-experiment, (18-b), the numeral *one* served to introduce the existential nominal. In the THERE sub-experiment, (18-c), we embedded the ∃ > ∀ configuration under existential *there*; given the definiteness restriction on existential *there*, this sub-experiment featured only ∃ > ∀ configurations (existential *there* refuses universally-quantified subjects; cf. Heim 1987). In the THEREONE sub-experiment, (18-d), the article *a* was replaced with the numeral *one*; as in the THERE sub-experiment, grammaticality permits only ∃ > ∀ configurations.

Each ORDER configuration occurred with either a surface or inverse INTERPRETATION-verifying image. A full PLAIN item appears in Figure 2. Again, the "inverse" image for the ∀ > ∃ configuration could in fact be verified by a surface interpretation of the sentence (Figure 2, bottom left).

3.1.3 Predictions

Given the well-documented availability for inverse scope in doubly-quantified sentences of English, we should find generally high ratings for inverse interpretations in our PLAIN sub-experiment. Due to the entailment patterns between interpretations, the critical test case for inverse scope is in the "inverse" condition of the "∃ > ∀" configuration; for a sentence to be judged true in this condition, participants must have given it an inverse parse. However, despite being rated as generally acceptable, the preference for surface

	∀ > ∃	∃ > ∀
SURFACE	"Every shark attacked a pirate."	"A shark attacked every pirate."
INVERSE	"Every shark attacked a pirate."	"A shark attacked every pirate."

Figure 2: Example item from Expt. 1 (Native English speakers tested on sentences of English).

interpretations should result in ratings for the critical condition that are on average lower than surface-scope conditions.

For the other sub-experiments, predictions are less clear. If the bi-clausal structure introduced by existential *there* precludes inverse scope, we should find significantly lower ratings—namely, ratings at floor—for inverse conditions in both the THERE and THEREONE sub-experiments. With the numeral *one*, we have no a priori reason to expect that its scope-taking abilities differ significantly from the article *a*, so ratings for the ONE sub-experiment should match those for the PLAIN sub-experiment.

3.1.4 Results

We split responses by sub-experiment; Figure 3 displays a violin plot of the raw ratings data, together with condition means and bootstrapped 95% confidence intervals drawn from 10,000 samples of the data (DiCiccio & Efron 1996). Recall that the THERE and THEREONE sub-experiments featured only the $\exists > \forall$ surface order. In what follows, we analyze the results of each sub-experiment in turn. For each analysis, we fit a mixed-effects ordinal regression model using the *ordinal* package (Christensen 2015) in *R*, predicting sentence ratings by scope INTERPRETATION ("surface" vs. "inverse") and TRIAL ORDER, and, in the PLAIN and ONE sub-experiments, by surface ORDER ("$\forall > \exists$" vs. "$\exists > \forall$").[7] Fixed effects predictors were centered before analysis. The models included random intercepts for participants and items.

Results from PLAIN sub-experiment. The model finds a main effect of ORDER ($\beta = -1.24$, $SE = 0.31$, $z = -3.93$, $p < 0.01$); $\exists > \forall$ configurations received lower ratings than $\forall > \exists$ configurations. The effect of INTERPRETATION was also significant ($\beta = -1.40$, $SE = 0.33$, $z = -4.23$, $p < 0.01$); "inverse" conditions received lower ratings than "surface". The interaction between ORDER and INTERPRETATION was not significant ($\beta = 1.00$, $SE = 0.62$, $z = 1.62$, $p < 0.11$), and neither was the effect of TRIAL ORDER ($\beta = -0.03$, $SE = 0.03$, $z = -1.22$, $p < 0.23$).
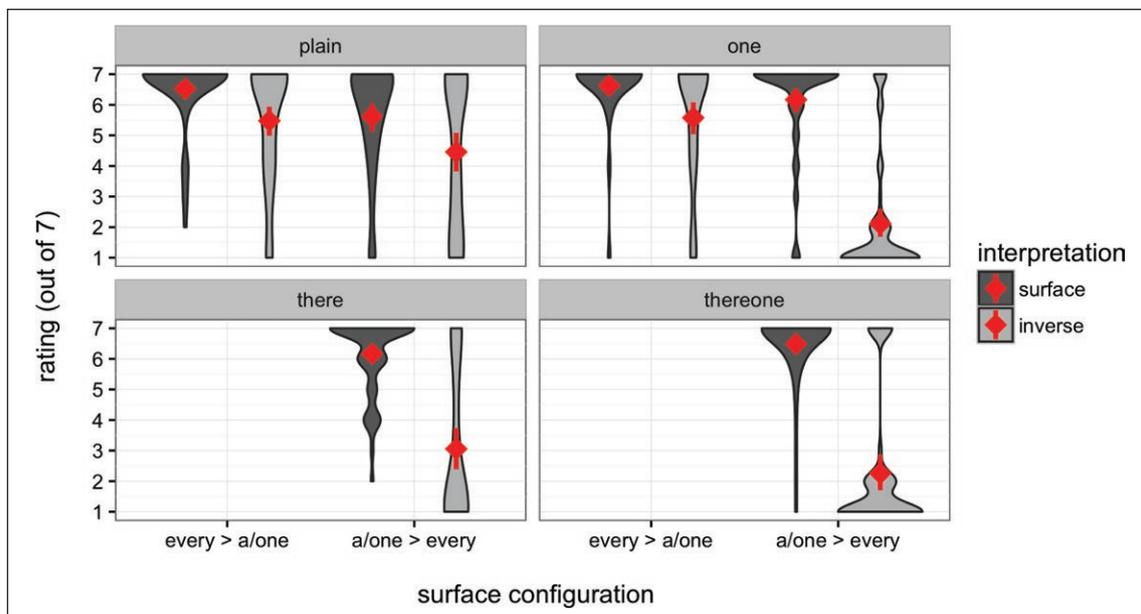

Figure 3: Expt. 1 results split by sub-experiment (native English speakers tested on English sentences).

---

[7] Performing mixed-effects linear regression analyses (Baayen et al. 2008) using the *lme4* package (Bates et al. 2014) in *R* yields the same pattern of results. The single exception is the effect of INTERPRETATION for the native English PLAIN sub-experiment, which did not reach significance in the linear regression analysis.

Results from ONE sub-experiment. The model finds a main effect of ORDER ($\beta = -1.99$, $SE = 0.32$, $z = -6.20$, p < 0.01); $\exists > \forall$ configurations received lower ratings than $\forall > \exists$ configurations. The model also finds a main effect of INTERPRETATION ($\beta = -2.61$, $SE = 0.34$, $z = -7.63$, p < 0.01); "inverse" conditions received lower ratings than "surface" conditions. Additionally, the model finds a significant interaction between ORDER and INTERPRETATION ($\beta = -2.05$, $SE = 0.60$, $z = -3.41$, p < 0.01); the critical inverse $\exists > \forall$ condition was rated much lower than each of the other three conditions. The effect of TRIAL ORDER was not significant ($\beta = -0.03$, $SE = 0.03$, $z = -1.06$, p < 0.29).

Results from THERE sub-experiment. Without an ORDER manipulation, we analyze only the effect of INTERPRETATION and TRIAL ORDER. The model finds a main effect of INTERPRETATION ($\beta = -2.51$, $SE = 0.42$, $z = -6.04$, p < 0.01); "inverse" conditions received lower ratings than "surface" conditions. The effect of TRIAL ORDER was not significant ($\beta = -0.00$, $SE = 0.03$, $z = -0.08$, p < 0.94).

Results from THEREONE sub-experiment. Again, we here analyze only the effect of INTERPRETATION and TRIAL ORDER. The model finds a main effect of INTERPRETATION ($\beta = -5.07$, $SE = 1.67$, $z = -3.04$, p < 0.01); "inverse" conditions received lower ratings than "surface" conditions. The effect of TRIAL ORDER was not significant ($\beta = 0.08$, $SE = 0.05$, $z = 1.58$, p < 0.12).

Comparing sub-experiments. Given our primary interest in the availability of inverse interpretations, we also compare responses to the critical inverse $\exists > \forall$ condition across sub-experiments. We therefore fit a model predicting sentence rating for just this condition by SUB-EXPERIMENT (together with TRIAL ORDER); the model included random intercepts for participants and items. Compared to the PLAIN baseline, the model finds significant effects of the ONE sub-experiment ($\beta = -2.74$, $SE = 0.49$, $z = -5.55$, p < 0.01), the THERE sub-experiment ($\beta = -1.82$, $SE = 0.48$, $z = -3.84$, p < 0.01), and the THEREONE sub-experiment ($\beta = -2.96$, $SE = 0.55$, $z = -5.33$, p < 0.01); ratings for the inverse $\exists > \forall$ condition were significantly higher when the construction featured indefinite *a* (instead of the numeral *one*) and mono-clausal syntax (instead of existential *there*). The effect of TRIAL ORDER was not significant ($\beta = -0.04$, $SE = 0.03$, $z = -1.41$, p < 0.16).

3.1.5 Discussion

Using the sentence-rating paradigm, we confirmed that English allows inverse scope in doubly-quantified sentences with indefinite *a* in subject position and universal *every* in object position. However, these inverse interpretations come at a cost, resulting in lower ratings for inverse vs. surface interpretations. Still, the average rating of 4.46 (out of 7) for inverse scope is completely in line with the work on English scope that precedes us. In general, complex structures are associated with lower ratings (see Gibson & Thomas 1999 for discussion), and the rating participants assign here signals that inverse scope is not impossible, but simply less likely than surface scope. This 4.46 acceptability rating, we claim, characterizes the availability of inverse scope. The effect of order, whereby $\exists > \forall$ configurations received lower ratings, is likely an artifact of our experimental design. Here we note the general dis-preference for indefinite subjects in transitive clauses, especially when the relevant sentences relate to pictures that flatten the event structure of an interpretation.

In addition to the PLAIN sub-experiment, we manipulated two properties of the English sentences—*a* vs. *one* and the presence of existential *there*—to yield three other sub-experiments. Ratings for the critical condition in each of the ONE (2.11), THERE (3.06), and THEREONE (2.26) sub-experiments were significantly lower than for the PLAIN sub-experiment, suggesting that adding the numeral *one* or bi-clausal *there* or both drastically reduces the availability of inverse scope. This result is expected for existential *there*, but

potentially surprising for the numeral *one*.[8] In what follows, we offer some thoughts on why this effect for *one* should not surprise us.

First, it should be noted that simple specificity-inference account of *one*'s behavior will not suffice to explain our data. Here is a sketch of such an account: the numeral *one* generally competes with *a* and engenders a specificity inference incompatible with inverse scope for $\exists > \forall$ sentences, that is, incompatible with a situation in which *one* corresponds to many (Figure 2, bottom right). But if *one* generally prefers to name just a single thing in a scene, we should find evidence of this preference also in $\forall > \exists$ configurations. In other words, we should find a preference for ostensibly *inverse* interpretations for such sentences, as they would allow *one* to name a single object (Figure 2, bottom left). However, we find the opposite: subjects reliably prefer surface-interpretation scenarios wherein *one* names multiple objects. Specificity writ broad will not do; *one* prefers to name a single referent only in subject position (i.e., in $\exists > \forall$ configurations).

Instead of triggering a general specificity inference, we suggest a processing explanation along the lines of the single-reference principle of Kurtzman & MacDonald (1993): listeners build an incremental parse of the sentences they hear; when they encounter *one* at the start of a sentence, they imagine just a single referent associated with it. This single-referent parse is incompatible with a one-as-many scenario (as in Figure 2, bottom right), accounting for participants' unwillingness to judge $\exists > \forall$ sentences with *one* as true in inverse, multi-referent scenarios. As we mentioned above, something like the single-reference principle is likely also active with indefinite *a* in our PLAIN sub-experiment, but its effect is less strong, presumably because *one* is phonologically more salient than *a*, so the pressure to build an initial, single-referent parse with *one* is more noticeable.

The question remains whether existential *there* or the numeral *one* stands to explain the availability of inverse scope in sentences of Mandarin. We return to this discussion in our analysis of the Mandarin.

3.2 Experiment 2: Native Mandarin

Having established the behavioral patterns from a language with the possibility for inverse scope (i.e., English) in our experimental paradigm, we now turn to Mandarin. As we mentioned at the outset, the status of inverse scope in Mandarin's doubly-quantified sentences has recently come into question. Our first task, then, is to resolve the debate on Mandarin's scope interpretation possibilities using the same paradigm from Expt. 1.

3.2.1 Participants

We recruited 132 participants (from either Mainland China or Taiwan) through a combination of email chains and advertisements on Chinese social media websites. 53 participants indicated that they were native speakers of Mandarin currently dominant in Mandarin; their data were included in the analysis presented below.[9]

3.2.2 Design, methods, and materials

We used a design similar to that used in Expt. 1, with the exception that sentences and instructions were presented in Mandarin, and participants could encounter an experimental condition at most two times. Sentences were translations of the English sentences from Expt. 1, adhering to the frames in (19).[10]

---

[8] See Tsai et al. (2014) for a fuller discussion of English existential *there* constructions in light of our ratings data.

[9] The exclusion rate appears high in these experiments because we recruited participants broadly via email chains and social media, but targeted a relatively narrow speaker profile.

[10] The full set of Mandarin test sentences appears in the Appendix.

(19)   *Sample sentences from Expt. 2*
   a.   ∀ > ∃ *sentence frame*:
         Mei-yi-tiao   shayu dou gongji-le   yi-ge   haidao.
         every-one-CLF   shark all attack-PST   one-CLF   pirate
         'Every shark attacked a/one pirate.'

   b.   ∃ > ∀ *sentence frame*:
         You   yi-tiao   shayu gongji-le   mei-yi-ge   haidao.
         exist one-CLF   shark attack-PST   every-one-CLF   pirate
         'A/one shark attacked every pirate.'

Sentences were recorded by an adult male native speaker of Mandarin from Beijing, and normed to ensure neutral intonation. We paired sentences with the same disambiguating pictures from Expt. 1, and asked participants to judge whether the sentence they heard appropriately described the picture using a seven-point Likert scale (1 = 'completely inappropriate', 7 = 'completely appropriate'). Participants completed a total of 15 trials (7 critical items and 8 fillers). Participants began the experiment with a short demographic questionnaire and three training slides to ensure that the audio played and the pictures were visible.

3.2.3 Predictions

If the literature on Mandarin scope stemming from Huang (1982) is correct in its conclusion that Mandarin does not allow inverse interpretations of doubly-quantified sentences, we should find ratings for the critical inverse ∃ > ∀ condition significantly lower than for all other conditions; in fact, these ratings should be at or near floor.

   If the pragmatics of the numeral *yi* or the contribution of existential *you* are sufficient to explain the prohibition on inverse scope in Mandarin, we should find that the ratings for inverse scope in Mandarin match the English ratings for the ONE, THERE, or THEREONE sub-experiments. However, if a factor beyond these elements (*yi* and *you*) is responsible for scope calculation in Mandarin, we should find that ratings for the Mandarin sentences are lower even than the low ratings for inverse scope in these English sub-experiments.

3.2.4 Results

Figure 4 plots the distribution of ratings, together with the average sentence ratings by condition with bootstrapped 95% confidence intervals drawn from 10,000 samples.

   We fit a mixed-effects ordinal regression model predicting sentence ratings by scope INTERPRETATION ("surface" vs. "inverse") and surface ORDER ("∀ > ∃" vs. "∃ > ∀"), as well as TRIAL ORDER. The model included random intercepts for participants and items. The model finds a main effect of INTERPRETATION ($\beta = -3.31$, $SE = 0.31$, $z = -10.65$,
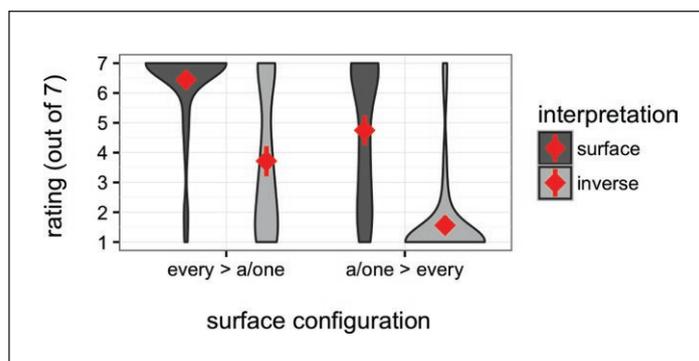

Figure 4: Expt. 2 results (native Mandarin speakers tested on Mandarin sentences).

p < 0.01) and a main effect of ORDER ($\beta = -2.50$, $SE = 0.29$, $z = -8.74$, p < 0.01). The interaction between INTERPRETATION and ORDER was not significant ($\beta = -0.41$, $SE = 0.51$, $z = -0.79$, p < 0.43). Inverse interpretations received lower ratings than did surface interpretations, and ∃ > ∀ configurations received lower ratings than did ∀ > ∃ configurations. As a result, the critical inverse ∃ > ∀ received the lowest ratings, but not lower than predicted by the addition of the two main effects. The effect of TRIAL ORDER was not significant ($\beta = 0.05$, $SE = 0.03$, $z = 1.40$, p < 0.17).

To compare the current results to the results from Expt. 1, we restricted our analysis to just the critical INVERSE ∃ > ∀ condition. We then fit a mixed-effects ordinal regression model predicting ratings to this condition by each sub-experiment (MANDARIN, PLAIN, ONE, THERE, THEREONE) and by TRIAL ORDER order. The model included random inter-cepts for participants and items. Compared to the MANDARIN baseline, ratings for the critical condition in each of the English sub-experiments were significantly higher (PLAIN: $\beta = 4.28$, $SE = 0.64$, $z = 6.64$, p < 0.01; ONE: $\beta = 1.67$, $SE = 0.53$, $z = 3.13$, p < 0.01; THERE: $\beta = 2.62$, $SE = 0.58$, $z = 4.53$, p < 0.01; THEREONE: $\beta = 1.49$, $SE = 0.55$, $z = 2.71$, p < 0.01). The effect of TRIAL ORDER was not significant ($\beta = -0.03$, $SE = 0.03$, $z = -1.07$, p < 0.29).

3.2.5 Discussion

Consistent with the consensus on inverse scope in Mandarin, participants demonstrated a strict resistance to inverse interpretations in our critical ∃ > ∀ configuration. In other words, our results support the hypothesis that Mandarin does not allow inverse scope in doubly-quantified sentences. This prohibition on inverse scope manifested as floor-level ratings, 1.56 out of a possible 7 points.[11]

This finding is at odds with the claim of Zhou & Gao (2009), namely that Mandarin does allow inverse scope in doubly-quantified sentences. As discussed above, the evidence Zhou & Gao use in support of their claim—judgments on ostensibly inverse interpreta-tions of ∀ > ∃ configurations—in fact cannot confirm the availability of inverse scope, given the entailment pattern between interpretations. Using instead ∃ > ∀ configurations which eschew the entailment problem, we have found strong evidence that Mandarin in fact does not allow inverse scope.

3.3 Comparing Mandarin and English

English allows inverse scope in doubly-quantified sentences; Mandarin does not.[12] Why these languages differ remains an open question. To better understand the lexico-syntactic properties of the Mandarin sentences that might account for this difference, we included in Expt. 1 two phenomena meant to more closely align the English and Mandarin com-parison: the numeral *one* (which could match the Mandarin indefinite *yi*) and existential *there* (which could approximate Mandarin *you*).

In English, substituting indefinite *a* with the numeral *one* yielded a marked decrease in ratings for inverse interpretations. Similarly, embedding the sentence under existential *there* drastically decreased inverse ratings. Including both *one* and *there* yielded the same decrease. However, despite receiving relatively low ratings, in each case English speak-ers consistently provided higher ratings for inverse interpretations than did Mandarin

---

[11] See Scontras et al. (2014) for a similar finding using a different, truth-value-judgment methodology. There it was observed that Chinese participants *never* judged the ∃ > ∀ configuration true in an inverse scenario.

[12] Throughout our investigation, we have been careful to limit our investigation to doubly-quantified sen-tences. See Tsai et al. (2014) for discussion of a broader set of Mandarin sentences, for example passive constructions and sentences with numerical expressions. Some of these more complex structures may actu-ally entail inverse readings, but their detailed investigation is still ahead.

speakers. Thus, while numeral semantics or bi-clausal syntax might contribute to the lack of inverse scope in Mandarin, alone they are unlikely to fully account for the prohibition. In fact, the stark unavailability of inverse scope in Mandarin suggests instead a language-wide ban such as the Isomorphic Principle (Huang 1982; Aoun & Li 1989).

4 Testing the heritage grammars

Now that we have confirmed that English and Mandarin are indeed quite different when it comes to possible scope ambiguities, we next explore the interaction between these two systems: What happens when one and the same individual presumably has access to both grammars? In other words, we test just how robust this ban on inverse scope would have to be by expanding our sights to the intersection of the English and Mandarin systems: heritage grammars.

4.1 Experiment 3: Heritage Mandarin on Mandarin

We start by investigating the heritage Mandarin grammar, testing English-dominant heritage speakers of Mandarin on sentences in Mandarin, their weaker language.

4.1.1 Participants

We recruited 140 participants through a combination of email chains and advertisements on Chinese language message boards. We identified as heritage speakers those participants who learned Mandarin as their first language, but were dominant in English and lived in the United States at the time of testing. Data from 26 heritage speakers of Mandarin were included in the analysis presented below.

4.1.2 Design, methods, and materials

The experiment was identical to Expt. 2, which tested native Mandarin speakers. However, all instructions were presented in English.

4.1.3 Predictions

If the prohibition on inverse interpretations of doubly-quantified sentences in Mandarin is robust to attrition and/or transfer from a dominant language, we should find that participants provide similarly low ratings to the critical inverse $\exists > \forall$ condition here. If, however, the costly and complex operation of calculating scope is susceptible to transfer, we should find a diverging pattern in the heritage grammar. Given the floor-level ratings for inverse scope in native Mandarin, this divergence could take but one direction: higher ratings for inverse scope in heritage Mandarin. If the availability of inverse scope transfers fully from English, we should find that heritage speakers provide ratings as high as those provided by native English speakers.

4.1.4 Results

Figure 5 plots the distribution of sentence ratings with averages by condition and bootstrapped 95% confidence intervals.

The analysis of sentence ratings was identical to that of Expt. 2. The model finds main effects of INTERPRETATION ($\beta = -2.48$, $SE = 0.39$, $z = -6.33$, p < 0.01) and of ORDER ($\beta = -1.97$, $SE = 0.38$, $z = -5.17$, p < 0.01); the interaction between INTERPRETATION and ORDER was not significant ($\beta = 0.85$, $SE = 0.74$, $z = 1.15$, p < 0.26). As in the native grammar, here we find that $\exists > \forall$ configurations are rated lower than $\forall > \exists$ and inverse interpretations are rated lower than surface, and that low ratings for the critical inverse $\exists > \forall$ are predicted by the addition of these two effects alone. The effect of TRIAL ORDER was not significant ($\beta = 0.02$, $SE = 0.04$, $z = 0.34$, p < 0.74).
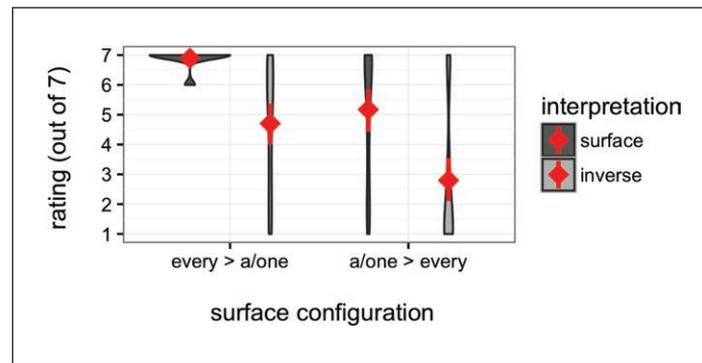
Figure 5: Expt. 3 results (English-dominant heritage speakers of Mandarin tested on Mandarin sentences).

Comparing heritage and native grammars. Next, we restrict our analysis to just the critical inverse $\exists > \forall$ condition and compare the current results to those of Expt. 2 with native Mandarin speakers. To do so, we fit a mixed-effects ordinal regression model predicting ratings to this condition by each population (MANDARIN vs. HERITAGE) and by TRIAL ORDER. The model included random intercepts for participants and items. Compared to the native baseline, English-dominant heritage speakers of Mandarin give significantly higher ratings to this critical condition ($\beta = 1.59$, $SE = 0.50$, $z = 3.16$, p < 0.01). The effect of TRIAL ORDER was not significant ($\beta = 0.03$, $SE = 0.06$, $z = 0.48$, p < 0.64).

We performed a similar comparison between the current results and those of the native English speakers from Expt. 1: we fit a mixed-effects ordinal regression model predicting ratings to the critical $\exists > \forall$ inverse condition by each population (ENGLISH vs. HERITAGE) and by TRIAL ORDER order. The model included random intercepts for participants and items. Here the model finds that the heritage Mandarin speakers' ratings are significantly *lower* than the comparable native English ratings of inverse scope in the PLAIN sub-experiment ($\beta = -1.44$, $SE = 0.45$, $z = -3.23$, p < 0.01). The effect of TRIAL ORDER was not significant ($\beta = -0.04$, $SE = 0.04$, $z = -1.03$, p < 0.31).

4.1.5 Discussion

In the comparison of heritage speakers with native speakers of both Mandarin and English, the picture that emerges suggests that these English-dominant heritage speakers of Mandarin do resist inverse interpretations for doubly-quantified sentences. Their ratings for the critical inverse condition were significantly lower than the other three conditions, and significantly lower than the English baseline for inverse scope (2.79 heritage Mandarin vs. 4.46 native English). However, heritage speakers' ratings were higher than the native Mandarin baseline (2.79 vs. 1.56 native Mandarin).

The higher ratings for inverse conditions (relative to native speakers) likely stem from a "yes-bias": heritage speakers are known to rate unacceptable or ungrammatical sequences higher than native controls (Benmamoun et al. 2013b; Laleko & Polinsky 2013; 2016; Orfitelli & Polinsky 2017). It is generally easier for heritage speakers to accept—rather than reject—linguistic material. In other words, when our heritage speakers heard a sentence that did not match the picture in the critical condition, they were less certain of this mismatch because they are less comfortable with their heritage grammar, and therefore they gave higher ratings than did the native speakers. Strongly supporting this yes-bias interpretation, a mixed-effects ordinal regression model predicting sentence ratings for all conditions by population (native Mandarin vs. heritage), with random

intercepts for participants and items, found that heritage speakers provided higher ratings across the board ($\beta = 0.53$, $SE = 0.20$, $z = 2.78$, $p < 0.01$).

Another possibility is that our heritage speakers actually find inverse interpretations in Mandarin more acceptable than do native speakers, owing to transfer from their dominant language, English. We have seen that English allows inverse scope, so perhaps this possibility has permeated the heritage Mandarin grammar to some degree.

In the next experiment, we attempt to resolve these competing hypotheses about the source of intermediate ratings for inverse scope in heritage Mandarin.

4.2 Experiment 4: Heritage Mandarin on English

Having found that heritage speakers of Mandarin resist inverse scope in Mandarin, although less severely than native speakers, we now shift the question to the source of these intermediate ratings: is the Mandarin grammar for scope in English-dominant heritage speakers experiencing the effects of transfer from an English grammar that does allow scope ambiguity? To address this question, we investigated the English grammar of scope in English-dominant heritage speakers of Mandarin.

4.2.1 Participants

We recruited 78 participants who did not participate in Expt. 3 through a combination of email chains and advertisements on social media and Chinese language message boards. We used the same criteria as in Expt. 3 to identify English-dominant heritage speakers of Mandarin. Data from 28 participants were included in the analyses presented below.

4.2.2 Design, methods, and materials

The experiment was identical to Expt. 1.

4.2.3 Predictions

If the English grammar of English-dominant heritage speakers of Mandarin is similar to the grammar of native English speakers, we should find similar patterns of ratings across our four sub-experiments when compared to the results of Expt. 1. Specifically, we should find low ratings for the critical inverse $\exists > \forall$ condition in each of the ONE, THERE, and THEREONE sub-experiments. We should also find relatively higher ratings for the inverse $\exists > \forall$ condition in the PLAIN sub-experiment, signaling that—like the native English speakers—English-dominant heritage speakers also allow inverse scope in doubly-quantified sentences.

If the English grammar of these heritage speakers differs from the native baseline, we might expect these participants to more closely align their patterns of ratings with the ratings of Mandarin sentences that we saw in Expts. 2 and 3. In other words, we should find a general resistance to inverse scope, regardless of the sub-experiment.

4.2.4 Results

Figure 6 plots the distribution of sentence ratings with condition means for each sub-experiment and bootstrapped 95% confidence intervals drawn from 10,000 samples of the data. Unless otherwise noted, data analysis was identical to that of Expt. 1.

Results from the PLAIN sub-experiment. The model finds a significant effect of ORDER ($\beta = -2.59$, $SE = 0.64$, $z = -4.04$, $p < 0.01$). The effect of INTERPRETATION was also significant ($\beta = -2.26$, $SE = 0.59$, $z = -3.84$, $p < 0.01$). The effect of TRIAL ORDER was not significant ($\beta = 0.02$, $SE = 0.05$, $z = 0.34$, $p < 0.74$), and neither was the interaction between ORDER and INTERPRETATION ($\beta = 0.03$, $SE = 1.15$, $z = 0.03$, $p < 0.98$).
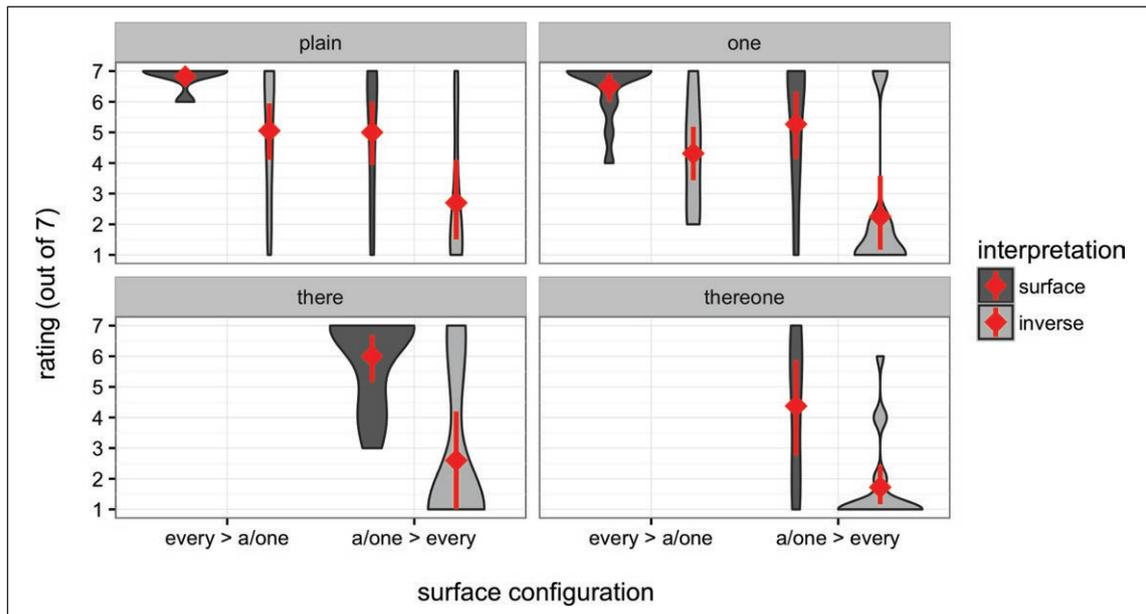
Figure 6: Expt. 4 results split by sub-experiment (English dominant heritage speakers of Mandarin tested on English sentences).

Results from the ONE sub-experiment. The model finds significant effects of ORDER ($\beta = -1.50$, $SE = 0.53$, $z = -2.84$, p < 0.01) and INTERPRETATION ($\beta = -2.36$, $SE = 0.50$, $z = -4.68$, p < 0.01). No other effects reached significance.

Results from the THERE sub-experiment.[13] The model finds a significant effect of INTERPRETATION ($\beta = -2.72$, $SE = 0.08$, $z = -2.88$, p < 0.01). No other effects reached significance.

Results from the THEREONE sub-experiment.[14] The model finds a significant effect of INTERPRETATION ($\beta = -3.76$, $SE = 1.20$, $z = -3.14$, p < 0.01). No other effects reached significance.

Comparing sub-experiments. Finally, we compared responses to the critical inverse $\exists > \forall$ condition across sub-experiments. Compared to the PLAIN baseline with indefinite *a* and no existential *there,* ratings to the critical condition in the other three sub-experiments did not differ significantly (ONE: $\beta = -0.49$, $SE = 1.12$, $z = -0.43$, p < 0.67; THERE: $\beta = -0.97$, $SE = 1.24$, $z = -0.78$, p < 0.44; THEREONE: $\beta = -1.94$, $SE = 1.24$, $z = -1.57$, p < 0.12). The effect of TRIAL ORDER was not significant ($\beta = -0.07$, $SE = 0.07$, $z = -1.00$, p < 0.32).

Comparing heritage and native grammars. To compare the current results with those from native speakers of English in Expt. 1, we again restricted our analyses to the critical inverse $\exists > \forall$ condition. We further restricted our attention to just the PLAIN sub-experiment for which we observed clear acceptance of inverse scope in the native speakers. We then fit a mixed-effects linear regression model predicting ratings to the inverse $\exists > \forall$ condition in the PLAIN sub-experiment by population (NATIVE VS. HERITAGE) and by TRIAL ORDER. The model finds that, compared to the NATIVE baseline, English-dominant heritage speakers of Mandarin provide significantly lower ratings to inverse scope ($\beta = -1.63$, $SE = 0.69$, $z = -2.38$, p < 0.05). The effect of TRIAL ORDER was not significant ($\beta = -0.04$, $SE = 0.05$, $z = -0.75$, p < 0.46).

---

[13] The ordinal regression model included only by-item random intercepts.
[14] The ordinal regression model included only by-item random intercepts.

4.2.5 Discussion

To evaluate the possibility of transfer of inverse scope from the English grammar in heritage speakers of Mandarin, we set out to establish whether this English grammar allows inverse scope in the first place. The results of the current experiment suggest that the English of these English-dominant heritage speakers of Mandarin does not allow inverse scope, or at least strongly resists it. These heritage speakers rated English inverse scope on average 2.55 out of a possible 7 points, nearly 2 full points below the 4.46/7 rating we observed in the native baseline. Given the observed lack of inverse scope in the English of English-dominant heritage speakers of Mandarin, it is unlikely that the intermediate ratings observed in Expt. 3 for heritage speakers tested in Mandarin stems from any transfer from a scope-allowing grammar. In fact, it would appear that these heritage speakers lack inverse scope in both their dominant English and their heritage Mandarin grammars. If anything, this fact may initially suggest transfer from Mandarin to English in our heritage speakers. However, the lack of inverse scope is more likely due not to Mandarin transfer, but to the strategy of adopting a more default (i.e., less encumbered) scope-calculation system. In what follows, we elaborate on this point.

4.3 The bigger picture

We found that heritage Mandarin speakers show a strong preference to avoid inverse scope (Expt. 3), suggesting a lack of transfer and a certain robustness to this prohibition in Mandarin. Still, the ratings that heritage speakers provided for Mandarin sentences were higher than the floor-level ratings that native speakers provided. Perhaps there is transfer from English, after all? Before settling on this conclusion, we decided to test the availability of inverse scope in the English of these heritage Mandarin speakers (Expt. 4). What we found was a lack of inverse scope, rendering less likely the possibility that a scope-shifting grammar (English) contaminated the rigid scope of Mandarin.It would seem, then, that the intermediate ratings observed in Expt. 3 for heritage Mandarin derive not from transfer, but from a lack of confidence on the part of heritage speakers in their weaker grammar, leading to the yes-bias in experimental settings. Taken together, our results suggest that the prohibition on inverse scope is a robust feature of Mandarin grammar that remains unchanged in the grammar of bilingual Mandarin speakers.

The question remains: what happened to the possibility of inverse scope in the English of these heritage speakers? In the heritage speakers we tested, even English resists inverse scope. Could it be that the lack of inverse scope transfers from Mandarin to English in our heritage speakers? Or might the relative expense of computing inverse scope, compounded with its reliance on a complex interaction between syntax, semantics, and pragmatics, render these interpretations too costly? We lack the data to settle this question once and for all, but there is one last population which might shed some light on its answer: heritage speakers of English dominant in a language that prohibits inverse scope.

Given the global status of English and the prevalence of English-speaking communities, tracking down heritage speakers of English is not a trivial task (see Viswanath 2013 for a discussion). Our target population is made more elusive by the requirement that these heritage speakers be dominant in a language that lacks inverse scope. We have so far tested four Japanese-dominant heritage speakers of English (i.e., bilinguals for whom English is the minority, home language) living in Japan.[15] Using the same English materials from Expts. 1 and 4, we observe that these heritage speakers rate the critical inverse

---

[15] Like Mandarin, Japanese appears to lack inverse scope in doubly-quantified sentences (Kuroda 1970; Kuno 1973; Hoji 1985; Han et al. 2008).

∃ > ∀ configuration an average of 2.13 out of a possible 7 points. Taking into account the 4.46/7 baseline observed for native English, it appears that these heritage English speakers equally lack inverse scope. Of course, these data are merely suggestive, but they do indicate that the trend may be in the direction away from inverse scope in English under contact.

To summarize: of the four populations (native vs. heritage; English vs. Mandarin) and five grammars (native English, heritage English, native Mandarin, heritage Mandarin, and the English of heritage Mandarin speakers), we find just one clear case of inverse scope: the native English grammar. We conclude in the following section with a discussion of why this might be so.

5 General discussion

Quantifier scope ambiguities feature prominently in many theories of the syntax-semantics interface, owing to the direct mapping from structure to meaning that generates the candidate readings. However, scope calculations are notoriously difficult, especially when the interpretation implicates a logical form seemingly at odds with the surface structure of the utterance. This is not surprising given that scope readings bring together at least three levels of representation: syntax, semantics, and pragmatics; facility with scope calculations presupposes a certain facility with each of these levels as well.

Preferences and dispreferences in scope interpretations are often accounted for under the notion of a pragmatic calculus (Musolino & Lidz 2006). Put simply, listeners start with the assumption that each interpretation is mapped to an unambiguous pattern, and only give up on that assumption if forced to do so in context. In other words, listeners assume a more economical model (one pattern : one interpretation) unless forced to map one pattern to more than one interpretation. This tendency toward economy often privileges surface interpretations, and helps to explain preferences like in English where inverse interpretations are possible though dispreferred. We replicated this pattern in Expt. 1 with our native English participants, who allowed inverse interpretations but gave higher ratings to surface ones.

In Mandarin, the picture looks rather different. We saw in Expt. 2 that native speakers of Mandarin resist inverse interpretations altogether (*pace* Zhou & Gao 2009). In contrast to the English dispreference, Mandarin appears to feature an all-out prohibition. Comparing English and Mandarin, we have some clues as to why. First, the predicate *you* 'exist' is generally obligatory with indefinite subjects in Mandarin. Second, Mandarin indefinites are headed by the numeral *yi* 'one'. We saw in Expt. 1 that both of these factors—existential structure and the presence of the numeral *one*—have a non-trivial effect on the availability of inverse scope in English. Perhaps *yi* and *you* have a similar effect in Mandarin. Still, Mandarin ratings for inverse interpretations were lower even than English sentences with *one* or an existential, leaving open the possibility for a grammaticalized ban such as the Isomorphic Principle (Huang 1982; Aoun & Li 1989).

Guided by previous work demonstrating the value of heritage language study to linguistic theory (Benmamoun et al. 2013a; 2013b; Scontras et al. 2015), we then investigated the robustness of the Mandarin prohibition on inverse scope in the context of potential transfer from a dominant English grammar that allows inverse scope. The results of Expt. 3 demonstrate a clear avoidance of inverse interpretations in English-dominant heritage speakers of Mandarin tested in Mandarin. These speakers come nowhere near the inverse-scope baseline observed for native English speakers, suggesting the lack of transfer from a scope-shifting English grammar. Moreover, when tested in English (Expt. 4), the speakers' dominant English grammar appears to lack the possibility for inverse interpretations in the first place.

But if heritage Mandarin speakers do not allow inverse scope, does it follow that they have a robust Mandarin grammar? Not necessarily. Heritage grammars are less dominant and more costly to employ. Heritage speakers might therefore prefer simpler grammars. A grammar with ambiguity will be more complex than one without it: such ambiguities require abandoning a one-to-one mapping between surface structures and interpretations. The heritage Mandarin speakers we tested might therefore be more likely to adopt the Mandarin-like system because it is simpler, in accordance with principles like *Processing Scope Economy* (Anderson 2004), which acknowledge the cost of inverse scope. In other words, a Mandarin-like grammar for scope is adopted by the heritage speakers not because it is inherited from the baseline, but because it happens to be simpler than the ambiguity-allowing alternative.

We find evidence for this line of reasoning in at least two additional domains: the English of our heritage Mandarin speakers, as well heritage English (i.e., the English of simultaneous and/or sequential bilinguals for whom English is the home (minority) language). Both grammars align with heritage Mandarin in their restriction on inverse interpretations for doubly-quantified sentences. Could it be that each of these groups lose the ability for inverse scope because the rigid scope grammar is simpler? In fact, this is precisely what the Lee et al. (2011) study found for English-dominant speakers with early exposure to Korean: a grammar lacking ambiguity. The confluence of evidence suggests that these bilinguals prefer less ambiguous grammars for scope—a preference visible in both the weaker and the dominant language. We fail to find interference from a dominant language when its system is more complex than the alternative. With this result in hand, future work should explore other domains of language where a decline in complexity takes precedence over effects from transfer.

Abbreviations

ACC = accusative, CLF = classifier, DECL = declarative, NOM = nominative, PROG = progressive, PST = past, RES = resultative

Additional File

The additional file for this article can be found as follows:

- **Appendix.** English and Mandarin test sentences. DOI: https://doi.org/10.5334/gjgl.198.s1

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

References

Anderson, Catherine. 2004. *The structure and real-time comprehension of quantifier scope ambiguity*. Evanston, IL: Northwestern University dissertation.
Aoun, Joseph & Y.-H. Audrey Li. 1989. Constituency and scope. *Linguistic Inquiry* 20. 141–172.

Aoun, Joseph & Y.-H. Audrey Li. 2003. *Essays on the representational and derivational nature of grammar.* Cambridge, MA: MIT Press.

Au, Terry Kit-Fong. 1983. Chinese and English counterfactuals: The Sapir- Whorf hypothesis revisited. *Cognition* 15. 155–187. DOI: https://doi.org/10.1016/0010-0277(83)90038-0

Baayen, R. H., D. J. Davidson & D. M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412. DOI: https://doi.org/10.1016/j.jml.2007.12.005

Bates, Douglas, Martin Maechler, Benjamin M. Bolker & Steven Walker. 2014. lme4: Linear mixes-effects models using Eigen and S4. ArXiv e-print; submitted to *Journal of Statistical Software;* http://arxiv.org/abs/1406.5823.

Becker, Michael & Jonathan Levine. 2010. Experigen: An online experiment platform. Available at https://github.com/tlozoot/experigen.

Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013a. Defining an "ideal" heritage speakers: Theoretical and methodological challenges. *Theoretical Linguistics* 39. 259–294. DOI: https://doi.org/10.1515/tl-2013-0018

Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013b. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39. 129–181. DOI: https://doi.org/10.1515/tl-2013-0009

Bergen, Leon & Noah D. Goodman. in prep. Quantifier domain restriction in probabilistic pragmatics.

Bochnak, Ryan & Lisa Matthewson (eds.). 2015. *Semantic fieldwork methodology.* Oxford: Oxford University Press.

Bruening, Benjamin. 2008. Quantification in Passamaquoddy. In Lisa Matthewson (ed.), *Quantification: A cross-linguistic perspective.* 67–103. Bingley: Emerald Press.

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

Christensen, Rune Haubo Bojesen. 2015. ordinal: Regression models for ordinal data. Available online at https://cran.r- project.org/web/packages/ordinal/.

Cooper, Robin. 1979. Variable binding and relative clauses. In F. Guen-thner & S. J. Schmidt (eds.), *Formal semantics and pragmatics for natural languages.* 131–169. Dordrecht: D. Reidel.

DiCiccio, Thomas J. & Bradley Efron. 1996. Boostrap confidence intervals (with discussion). *Statistical Science* 11. 189–228. DOI: https://doi.org/10.1214/ss/1032280214

Fodor, Janet Dean. 1982. The mental representation of quantifiers. In Stanley Peters & Esa Saarinen (eds.), *Processes, beliefs, and questions.* 129–164. Dordrecht: Reidel. DOI: https://doi.org/10.1007/978-94-015-7668-0_5

Gibson, Edward & James Thomas. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14(3). 225–248. DOI: https://doi.org/10.1080/016909699386293

Grodner, Daniel J. & Edward A. F. Gibson. 2005. Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science* 29. 261–290. DOI: https://doi.org/10.1207/s15516709cog0000_7

Han, Chung-hye, Dennis Ryan Storoshenko & Yasuko Sakurai. 2008. An experimental investigation into the placement of the verb in the clause structure of Japanese. In *Current issues in unity and diversity of languages: Collection of the papers selected from the 18th International Congress of Linguistics.* Seoul: The Linguistic Society of Korea.

Han, Chung-hye, Jeffrey Lidz & Julien Musolino. 2007. V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* 38. 1–48.

Heim, Irene. 1987. Where does the Definiteness Restriction apply? Evidence from the definiteness of variables. In Eric J. Reuland & Alice G. B. ter Meulen (eds.), *The representation of (in)definiteness.* 21–42. Cambridge, MA: MIT Press.

Hoji, Hajime. 1985. *Logical form constraints and configurational structures in japanese.* Seattle, WA: University of Washington dissertation.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of gram- mar*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Huang, C.-T. James. 1987. Existential sentences in Chinese and (in)definiteness. In Eric J. Reuland & Alice G. B. ter Meulen (eds.), *The representation of (in)definiteness.* 226–253. Cambridge, MA: MIT Press.

Huang, Shuan-Fan. 1981. On the scope phenomena of Chinese quantifiers. *Journal of Chinese Linguistics* 9. 226–243.

Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide scope indefinites? In Susan Rothstein (ed.), *Events and grammar.* 163–196. Dordrecht: Kluwer.

Kuno, Susumu. 1973. *The structure of the Japanese language.* Cambridge, MA: MIT Press.

Kupisch, Tanja. 2013. A new term for a better distinction? a view from the higher end of the proficiency scale. *Theoretical Linguistics* 39. 203–214. DOI: https://doi.org/10.1515/tl-2013-0012

Kupisch, Tanja & Jason Rothman. 2016. Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism.* DOI: https://doi.org/10.1177/1367006916654355

Kuroda, Shige-Yuki. 1970. *Japanese syntax and semantics.* Dordrecht: Kluwer.

Kurtzman, Howard S. & Maryellen C. MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48. 243–279. DOI: https://doi.org/10.1016/0010-0277(93)90042-T

Laleko, Oksana & Maria Polinsky. 2013. Marking topic or marking case: A comparative investigation of heritage Japanese and heritage Korean. *Heritage Language Journal* 10. 40–64.

Laleko, Oksana & Maria Polinsky. 2016. Between syntax and discourse: Topic and case marking in heritage speakers and L2 learners of Japanese and Korean. *Linguistic Approaches to Bilingualism* 6. 396–439. DOI: https://doi.org/10.1075/lab.14018.lal

Lee, H.-T. Thomas. 1986. *Studies on quantification in Chinese.* Los Angeles, CA: University of California, Los Angeles dissertation.

Lee, Miseon, Hye-Young Kwak, Sunyoung Lee & William O'Grady. 2011. Processing, pragmatics, and scope in Korean and English. In Ho-Min Sohn, Haruko Minegishi Cook, William O'Grady, Leon Serafim & Sany Yee Cheon (eds.), *Japanese/Korean linguistics,* vol. 19. Stanford, CA: CSLI Publications.

Lee, Sunyoung. 2009. *Interpreting scope ambiguity in first and second language processing: Universal quantifiers and negation.* Honolulu, HI: University of Hawai'i at Manoa dissertation.

Liu, Yi-Hsien. 2011. Topic-comment and the Chinese existential construction. In Zhou Jing-Schmidt (ed.), *Proceedings of the 23rd North American Conference on Chinese Linguistics*, vol. 2. 48–65. Eugene: University of Oregon.

Lohndal, Terje. 2013. Generative grammar and language mixing. *Theoretical Linguistics* 39. 215–224. DOI: https://doi.org/10.1515/tl-2013-0013

May, Robert. 1977. *The grammar of quantification.* Cambridge, MA: Massachusetts Institute of Technology dissertation.

May, Robert. 1985. *Logical form: Its structure and derivation.* Cambridge, MA: MIT Press.

Mayr, Clemens & Benjamin Spector. 2011. Not too strong! Generalizing the Scope Economy condition. In Martin Prinzhorn, Viola Schmitt & Sarah Zobel (eds.), *Proceedings of Sinn und Bedeutung* 14. 305–321.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In K. J. J. Hintikka, J. M. E. Moravcsik & P. Suppes (eds.), *Approaches to natural language,* vol. 49 Synthese Library. 221–242. Dordrecht: Reidel.

Montrul, Silvina. 2016. *The acquisition of heritage languages.* Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139030502

Musolino, Julien & Jeffrey Lidz. 2006. Why children aren't universally successful with quantification. *Linguistics* 44. 817–852. DOI: https://doi.org/10.1515/LING.2006.026

O'Grady, William, Miseon Lee & Hye-Young Kwak. 2009. Emergentism and second language acquisition. In William C. Ritchie & Tej K. Bhatia (eds.), *Handbook of second language acquisition.* 68–81. Bingley, UK: Emerald Press.

Orfitelli, Robyn & Maria Polinsky. 2017. When performance masquerades as comprehension: Grammaticality judgments in non-native speakers. In Mikhail Kopotev, Arto Mustajoki & Olga Lyashevskaya (eds.), *Quantitative approaches to the Russian language.* London: Routledge.

Pascual y Cabo, Diego, Anne Lingwall & Jason Rothman. 2012. Applying the Interface Hypothesis to heritage speaker acquisition: Evidence from Spanish mood. In Alia K. Biller, Esther Y. Chung & Amelia E. Kimball (eds.), *Proceedings of the 36th annual Boston University Conference on Language Development.* 437–448. Somerville, MA: Cascadilla Press.

Pillot, Carl, Gregory Scontras & Lauren Clemens. 2012. ExperigenRT: Measure reaction times in web-based auditory experiments. Available online at http://pollab.fas.harvard.edu/?q=node/212.

Polinsky, Maria. 2011. Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition* 33. 305–328. DOI: https://doi.org/10.1017/S027226311000077X

Quine, Willard Van Orman. 1960. *Word and object.* Cambridge, MA: MIT Press.

Reinhart, Tanya. 1976. *The syntactic domain of anaphora.* Cambridge, MA: Massachusetts Institute of Technology dissertation.

Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20. 335–397. DOI: https://doi.org/10.1023/A:1005349801431

Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13(2). 155–163. DOI: https://doi.org/10.1177/1367006909339814

Ruys, E. G. 1992. *The scope of indefinites.* Utrecht: Utrecht University dissertation.

Scontras, Gregory, Cheng-Yu Edwin Tsai, Kenneth Mai & Maria Polinsky. 2014. Chinese scope: An experimental investigation. In Urtzi Etxeberria, Anamaria Fălăuş, Aritz Irurtzun & Bryan Leferman (eds.), *Proceedings of Sinn und Bedeutung* 18. 396–414.

Scontras, Gregory, Zuzanna Fuchs & Maria Polinsky. 2015. Heritage language and linguistic theory. *Frontiers in Psychology* 6. 1545. DOI: https://doi.org/10.3389/fpsyg.2015.01545

Sorace, Antonella. 2011. Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism* 1. 1–33. DOI: https://doi.org/10.1075/lab.1.1.01sor

Tsai, Cheng-Yu Edwin, Gregory Scontras, Kenneth Mai & Maria Polinsky. 2014. Prohibiting inverse scope: An experimental study of Chinese vs. English. In Christopher Piñon (ed.), *Empirical issues in syntax and semantics* 10. 305–322.

Viswanath, Arun. 2013. Heritage English in Israeli children. B.A. thesis. Harvard University.

Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20. 399–467. DOI: https://doi.org/10.1023/A:1005354323136

Yeh, David & Dedre Gentner. 2005. Reasoning counterfactually in Chinese: Picking up the pieces. In *Proceedings of the twenty-seventh annual meeting of the Cognitive Science Society*. 2410–2415.

Zhou, Peng & Liqun Gao. 2009. Scope processing in Chinese. *Journal of Psycholinguistic Research* 38. 11–24. DOI: https://doi.org/10.1007/s10936-008-9079-x