RESEARCH

In support of representational economy: Agreement in heritage Spanish

Gregory Scontras[1], Maria Polinsky[2] and Zuzanna Fuchs[3]
[1] University of California, Irvine, US
[2] University of Maryland, US
[3] Harvard University, US
Corresponding author: Gregory Scontras (g.scontras@uci.edu)

This paper investigates the morphosyntax of number and gender agreement in English-dominant heritage speakers of Spanish. Our study builds on the experimental paradigm of agreement attraction developed by Fuchs et al. (2015) and elicits responses to agreement failures to diagnose the potentially-independent contributions of number and gender features. By comparing the structuring of these agreement categories in native and heritage Spanish, we document a case of grammatical divergence: heritage speakers restructure their agreement categories, favoring fewer feature values and less structure. The principled nature of this restructuring demonstrates that the loss of agreement in heritage languages is not superficial or accidental, but rather follows from predictable structural changes.

1 Introduction

One of the main goals of linguistic theory is to formalize what it is we know when we know a language. As the language varies, so too does the knowledge that underlies it. This observation is obvious to the point of being trivial: speakers of Spanish know Spanish, and their knowledge differs from that of speakers of English, who know English; hence the difference in language. As linguists explore various constraints on the variation that is allowed across languages, cross-linguistic variation leads them to the principles underlying Universal Grammar. Here we address a potentially different but *prima facie* no less informative type of variation, that between speakers of the *same* language. One might call this variation "intra-linguistic."

We take as our starting point the observation that there are many different types of speakers. There are monolingual, literate, educated native speakers who know their language perfectly—the Chomskyan ideal stated in the well-known quote: "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). There are L1 learners, often monolingual as well: children on their way to fully acquiring the language. There are also L2 learners, (young) adults who are actively learning (or have actively learned) another language in addition to their first. From the perspective of traditional linguistic theory, these count as "typical"

speaker profiles. For our purposes, the real excitement centers around "atypical" cases where the normative developmental trajectory does not attain; the speakers that result at once fit all and none of the profiles just described. We have in mind *heritage speakers*: (relatively) unbalanced bilinguals who shifted from their first language (their heritage language) to their dominant language early in childhood.[1]

According to most definitions, heritage speakers are individuals who were raised in homes where a language other than the dominant community language was spoken, resulting in some degree of bilingualism in both the heritage language and the dominant language (Valdés 2000; Rothman 2009; Benmamoun et al. 2013a; b; Scontras et al. 2015; Montrul 2016). This relatively unconstrained definition makes it almost impossible to give a concrete model for a heritage speaker, but this is by design: heritage language proficiency falls along a continuum. On one end, heritage speakers can be so proficient that in casual conversation they pass as native speakers. On the other end, some heritage speakers are able to comprehend but unable to fluently speak their heritage language; these are the so-called "overhearers" (Au et al. 2008; Chang 2016). The variability among heritage speakers derives from the many pathways that lead to the attainment of language under bilingualism, with significant variation in the amount and quality of input.

The notion of heritage speakers typically subsumes simultaneous or early sequential bilinguals. Some heritage speakers are born to immigrant parents in a country where the heritage language is not the native tongue; these bilinguals may be exposed to both the heritage language and the dominant language of the broader community from the outset of acquisition. Other heritage speakers are born in the home country of their parents, and for the early years of their life encounter only the heritage language. They often begin acquiring their second, dominant language at the time of emigration, when they shift to a different speech community. Within sequential bilinguals, the age of emigration or first exposure to the dominant language affects the resulting language proficiency (for extensive discussion, see Montrul 2008; 2016). Still more variation arises from the amount of input and maintenance speakers receive for the heritage grammar: some speakers use the heritage language at home regularly with their families, while others have only infrequent opportunities to perform in their heritage language.[2]

Despite the heterogeneity of their demographic profiles and levels of proficiency, certain characteristics of heritage speakers have been shown to be remarkably consistent, setting these speakers apart both from native and L2 speakers of a given language. At the phonological level, heritage speakers diverge from their native counterparts in aspects of pronunciation and prosody (Godson 2003; Barlow 2014; Chang 2016). At the syntactic level, heritage speakers tend to impose more rigid word order where native speakers allow for flexibility (Isurin & Ivanova-Sullivan 2008; Ivanova-Sullivan 2014); relatedly, they limit their inventory of syntactic dependencies (Polinsky 2011). At the intersection of syntax, semantics, and pragmatics, evidence from Chinese and English heritage speakers suggests that heritage grammars lack quantifier scope ambiguities (Scontras et al. 2017). The current investigation follows up on findings at the morphological level, where heritage speakers are known to eliminate irregular forms and struggle with inflectional morphology (Benmamoun et al. 2013a; b).

---

[1] We use the terms "native" and "heritage" in recognition of the documented differences between these two groups, not to imply that heritage speakers are not "native" in their own right. We return to this point in Section 5.

[2] The dominant language of a speech community need not always be the heritage speaker's stronger language (for discussion of specific cases, see Meisel 2007; Kupisch 2012). For the purposes of our study, the dominant language and the stronger language happen to coincide, with the heritage language being the weaker one.

In the nominal domain, heritage speakers are known to perform differently from native speakers in the production and comprehension of agreement in gender (see Håkansson 1995 for Swedish; Montrul et al. 2008 for Spanish; Polinsky 2008a for Russian), marking of definiteness (Håkansson 1995 for Swedish; Bolonyai 2016 for Hungarian), case-marking (Polinsky 1995; 2006; 2008a; b for Russian; Song et al. 1997 for Korean), and topic-marking (Laleko & Polinsky 2013; 2016 for Japanese and Korean). Despite the vulnerability of these categories—many of which also pose a challenge to L1 and L2 learners—there are also linguistic domains in which heritage speakers demonstrate remarkably resilient knowledge, for example, tense (even in recessive bilinguals, as shown by work on Inuttitut; see Sherkina-Lieber et al. 2011) or numerical expressions (Polinsky 2016a).

The question thus arises as to precisely *what* heritage speakers know about their heritage morphology. The problem lies in using observed differences in linguistic *behavior* to diagnose and distinguish differences in linguistic *knowledge*. The null hypothesis holds that heritage speakers possess the same knowledge as native speakers: a heritage speaker of Spanish has the same grammatical knowledge of Spanish as a native speaker of Spanish. If the knowledge is the same, then any differences we observe in behavior arise from differences in the usage of that knowledge. Faced with the costly task of performing in their less dominant language, heritage speakers, even the most proficient ones, would diverge from native speakers because of the overwhelming processing load. Under the null hypothesis, observed differences are symptomatic not of a diverging heritage grammar, but of the relative scarcity of online resources.

The competing hypothesis, which we investigate experimentally below, holds that differences in performance are symptomatic of deeper, structural differences in the heritage grammar. Finding evidence that points to a true divergence between native and heritage grammars, one must then confront the question of what pressures led to the difference. To answer this question, we must arrive at a clear characterization of the structural difference: does the heritage grammar reduce or augment structure relative to the native baseline? The pressures driving each outcome are interestingly different. Finding reduced structure relative to the native baseline, heritage speakers will have likely prioritized representational economy, restructuring their grammar in favor of lighter-weight linguistic representations. Less articulated, more parsimonious structures (e.g., structures with fewer explicit agreement features or syntactic projections) could ease the load on working memory and might therefore be preferred to their fully-articulated brethren. Pressure from representational economy would therefore lead to a general slimming-down of the heritage morphology.

Conversely, heritage speakers may prioritize analyticity, restructuring native grammars with more-fully-articulated syntactic structures. Increasing analyticity with syntactic structures that allow one-to-one correspondences between underlying features and surface forms would lead to a richer, more rigid morphology (Sorace 2005; Keating et al. 2016; Montrul 2016). The research to date suggests that increased analyticity is commonly the trigger for restructuring in heritage grammar. It bears noting that we use the terms "representational economy" and "increased analyticity" without advancing any specific agenda on theories of sentence processing (for example, theories of "shallow" vs. "full" parsing; Ferreira et al. 2002; Sanford & Sturt 2002; Clahson & Felser 2006). Our intent is merely to operationalize these notions in such a way as to deliver diverging predictions with respect to levels of feature articulation: a less-well-articulated feature space (i.e., fewer feature values and less structure) would evidence pressures from representational economy, while increased analyticity predicts fully-specified feature values and fully-articulated structure.

To investigate the content of heritage speakers' knowledge about their morphology, together with the pressures that could shape differences in that knowledge relative to the native baseline, here we extend the experimental methodology developed in Fuchs et al. (2015) to examine number and gender agreement in heritage Spanish. We focus on number and gender features because they are expressed on independent exponents in the native Spanish grammar, and because theoretical approaches to the structure of these categories provide two analytical possibilities for their representation in syntax. By comparing the structuring of these agreement categories in native and heritage Spanish, we aim to demonstrate that the loss of agreement in heritage languages is not superficial or accidental, but rather follows from principled and predictable structural changes.

2 Background: Feature representation in the native Spanish baseline

Grammatical agreement happens when features of one constituent (e.g., a sentential predicate) purposefully align with the features of another (e.g., a sentential subject). The specific form that a constituent expresses is determined by the features that the constituent holds (e.g., *I* vs. *they*; *runs* vs. *run*). Languages vary in the quality and quantity of agreement they express; the recurrent feature categories that participate in agreement are person (e.g., first vs. second), number (e.g., singular vs. plural), and gender (e.g., feminine vs. masculine). Once we assume that agreement features are represented structurally and merged into a syntactic tree, the question shifts to precisely how those structures are built.

Explorations of *feature geometry* investigate the structural organization and hierarchical relationships in agreement systems (Ritter 1993; Harley 1994; Harley & Ritter 2002). Feature geometry identifies the feature categories of person, number, and gender; here we limit our investigation to number and gender only. This narrowing of focus is motivated in two ways. First, person stands apart from the other features in its distribution (for example, unlike the other features, person agreement never appears on adjectives; see Baker 2008). Second, the hierarchical position of person relative to the other feature categories has been well established: it dominates the other features (Harley & Ritter 2002; among others). Meanwhile, the relationship between gender and number is less clear.

Under the assumption that both number and gender are represented in the syntax, there are two competing models of their structural relationship.[3] One model analyzes the two features as *bundled* together (cf. Ritter 1993; Carstens 2000), hosted on the head of the same projection:

(1) $[_{NumP} \ldots [\text{Num}_{\{Number, Gender\}}]]$

Under this bundling model, gender and number features are projected together, which means they are dependent on each other (Antón-Méndez et al. 2002). The effect is that valuation of one feature presupposes valuation of the other; number and gender agreement is part of a single process whereby one head probes for both features at the same time.

---

[3] In fact, there is a third model for the syntactic representation of gender, known as distributed gender, in which gender features in a single language occur in more than one locus on the nominal spine. However, motivation for such a model comes strictly from considerations of animate/interpretable gender. This model is commonly implemented as providing one locus for biological sex/animate gender and a separate locus for grammatical gender, that one lower than the animate gender projection. Because animate gender is beyond the scope of this paper, we will not address this model here. For discussion see Picallo (2007; 2008); Kramer (2009; 2016); Steriopolo & Wiltschko (2010); Armoskaite (2011); Matushansky (2013); Atkinson (2015); Fathi & Lowenstamm (2016).

Empirical support for the bundling model begins with the observation that languages regularly combine number and gender, fusing the two features into a single morpheme that expresses changes in one or both features. Few languages have systems in which both number and gender features participate in agreement but remain completely independent of each other, and it is hard to find a language that reliably this description; at the writing of this paper, candidate languages include Abkhaz (Hewitt 1979), Amharic (Kramer 2015: 44), Yimas (Foley 1991), and Romanian, at least according to some researchers (Giurgea 2008; Croitor & Giurgea 2009).

The alternative to the bundling model *splits* number and gender features; they are projected and valued independently (Picallo 1991; Antón-Méndez et al. 2002; Carminati 2005). Under the split model, gender morphology hosted on the nominal spine heads its own gender projection (GenP), which is dominated by the number projection (NumP), as in (2).[4] Thus, number and gender agreement are valued separately, independently of each other.

(2) $[_{\text{NumP}} \cdots [_{\text{GenP}} \cdots$

Evidence for the split model comes from the observation that number and gender morphology is consistently linearly ordered. In those languages, including Spanish, where the two features can be pre-theoretically separated into independent morphemes, the morpheme order is Stem-Gender-Number:

(3)   a. $[[\textit{libr}]\text{-}[_{\text{GenP}} o]\text{-}[_{\text{NumP}} s]]$ 'books'
         b. $[[\textit{libr}]\text{-}[_{\text{GenP}} o]\text{-}[_{\text{NumP}} ø]]$ 'book'

Under the split model—with the structure shown in (2) above—the morpheme order of Stem-Gender-Number is a direct consequence of the fact that number hierarchically dominates gender. Under a bundling model, number and gender are bundled together and therefore leveled—there is no clear hierarchical relationship between the features that could deliver a consistent linear order.

With arguments in favor of both the bundled model and the split model, the task is to determine which model is appropriate for which language—it need not be the case that every language approaches number and gender features in the same way (for a discussion of yet more arguments, see Fuchs et al. 2015 and the references therein). In Fuchs et al. (2015), we used experimental evidence from agreement phenomena to establish the baseline feature representation and specification in the native Spanish grammar. Our assumption, then and now, is that grammar and its parser are in an isomorphic relation, so that observing the parser allows for observation of the grammar (Phillips 2013). The work in this approach centers around developing testable hypotheses about *behavior* on the basis of articulated theories of *grammar.* Our strategy was to elicit responses to agreement failures in a way that would diagnose the potentially independent contributions of number and gender features. Before presenting the details of our experiment, we will present a quick overview of the relevant facts concerning Spanish agreement.

---

[4] In this schema, we show two phrasal nodes, NumP and GenP, in the structure of a nominal. However, many researchers represent gender as a feature on the nominal stem, thus present on the categorizer head n (cf. Alexiadou 2004; Kramer 2015; 2016). We will return to the latter representation in Section 4.3 below.

Agreement in number and gender in Spanish is pervasive: determiners, adjectives, and participles must all agree with a head noun in both number and gender. The Spanish plural is formed by suffixing *-s* onto a noun, while the singular is understood as the absence of the plural morpheme (in other words, it appears morphologically bare).[5]

(4)   a.   la   manzana
                the.F.SG apple.F.SG

          b.   el   libro
                the.M.SG book.M.SG

          c.   las   manzanas
                the.F.PL apple.F.PL

          d.   los   libros
                the.M.PL books.M.PL

Our focus here is on morphologically-transparent inanimate nouns with uninterpretable gender (as opposed to animate and human nouns whose gender is often interpretable and may be encoded separately, represented higher in the structure of a nominal; see footnote 3). Spanish has two surface genders: masculine and feminine; they are relatively equally distributed in the nominal lexicon (Bull 1965). The two genders also appear to be equally specified morphologically: the theme vowel *-o* is most often associated with the masculine gender, and the theme vowel *-a* is most often associated with the feminine gender.[6]

To summarize: the presence (or absence) of a word-final *-s* provides cues to that word's number, and the vowel (*-a-* vs. *-o-*) that would precede *-s* provides cues to the word's gender. As the number- and gender-agreement morphemes are in principle morphologically independent, we could manipulate their combination to produce sentences with different kinds of agreement errors. Because the bundling and split models of feature representation make different commitments regarding the valuation of agreement features, the predictions of the two models pull apart in cases where one or both features are incorrectly matched. These are cases of agreement attraction (Bock & Miller 1991; Bock & Eberhard 1993).

In agreement attraction, errors in agreement are masked by the features of a distractor constituent. For subject-verb agreement, the predicate fails to agree with its subject, but rather agrees with the features of some other noun present in the same clause (typically, but not necessarily, closer to the agreeing verb in the surface order). The paradigm example appears in (5), where the predicate (shown in SMALL CAPS) should be singular to match the singular feature of the head noun (goal, shown in **bold**), but instead it is plural, matching the plural feature of the so-called "local" noun (distractor, *italicized*). The agreement error yields an ungrammatical sentence, yet such cases are regularly produced and perceived as grammatical.

---

[5] In some dialects of Latin American Spanish, particularly in Caribbean Spanish and coastal varieties of Argentinian and Chilean Spanish, *-s* in coda position undergoes weakening (for an overview, see Lipski 1984; 2011; Guitart 1997). This weakening should have a minimal effect on the results we present below given the demographics of our participants, together with the exposure heritage Spanish speakers receive from speakers of various dialects.

[6] The assumption that *-o* and *-a* are the morphological realizations of masculine and feminine gender, respectively, has been challenged, most notably by Harris (1991). Still, these exponents serve as strong cues allowing the parser to predict gender.

(5)   The **key** to the *cabinets* WERE lost.

Agreement attraction has been widely studied in behavioral experiments, and a substantial body of literature has identified factors that affect the probability that a case of attraction will be perceived as grammatical. Fuchs et al. (2015) put agreement attraction to use in exploring the difference between bundling and split approaches to feature representation. If number and gender features are bundled together, they ought to be valued simultaneously. Thus, the number and gender features of a noun (more accurately, the single feature bundle of that noun) should determine agreement together, at the same time. With simultaneous feature valuation, when an incorrect noun enters into agreement with an adjective, both its number and gender features should effect agreement attraction. To illustrate this point, consider the following ungrammatical sentences of Spanish:

(6)   a. *El   niño considera   la   noticia   en las
            the.M.SG boy   consider.PRS.3SG the.F.SG news.item.F.SG in the.F.PL
            revistas   terriblemente aburridas.
            magazine.F.PL terribly   boring.F.PL
            Intended: 'The boy considers the news item in the magazines to be terribly boring.'

        b. *El   niño considera   la   noticia   en los
            the.M.SG boy   consider.PRS.3SG the.F.SG news.item.F.SG in the.M.PL
            periódicos   terriblemente aburridos.
            magazine.M.PL terribly   boring.M.PL
            Intended: 'The boy considers the news item in the magazines to be terribly boring.'
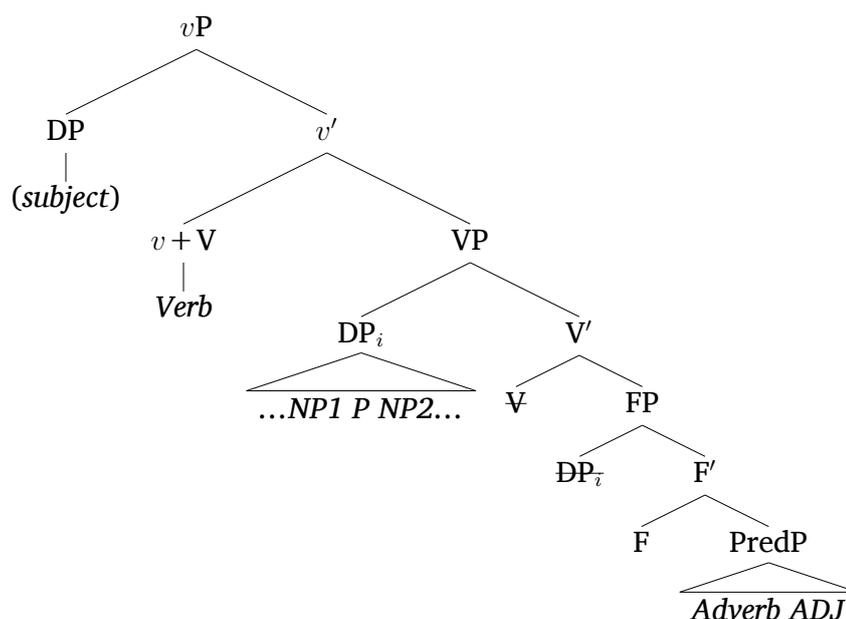
In each sentence, the local noun (i.e., *revistas* or *periódicos*) has spread its features to the predicate (i.e., *aburridos/-as*), which may lead to the illusion of grammaticality.[7] If number and gender are projected and valued together, as per bundling approaches, then when the predicate (erroneously) gets a feature (e.g., number) from the local noun, it should get the other feature (e.g., gender) as well. In other words, agreement attraction in one feature ought to precipitate agreement attraction in the other feature, with the result that both of the above sentences should be rated equally high (or equally low).

If number and gender are split, then the features are projected and valued independently, and agreement attraction in number can proceed independently of attraction in gender. The violations are evaluated (or ignored) on their own merits. If attraction in number does not carry the local noun's gender features along for the ride, then the sentence in (6a)—where the predicate agrees with the local noun in number but matches the head noun in gender—should be rated higher than (6b), where neither feature of the head noun correctly appears on the predicate. It was precisely these sorts of sentences that we set out to test.

Fuchs et al. (2015) designed an auditory sentence-rating task in which participants rated the acceptability of sentences they heard. Each stimulus was based on the syntactic frame in (7).

---

[7] Note that the sentences in (6) would be grammatical if the predicate modified the local noun. To ensure that this unintended parse is not available, Fuchs et al. (2015) conducted a norming study, which demonstrated the low probability of this parse.

(7)   (Subject) Verb NP1 PP NP2 Adverb Adjective/Participle

```
                        vP
                   ┌────┴────┐
                  DP         v′
                   │      ┌───┴───┐
              (subject)  v + V       VP
                           │      ┌───┴────┐
                         Verb    DPᵢ        V′
                               ┌──┴──┐   ┌──┴──┐
                            …NP1 P NP2… V      FP
                                            ┌──┴──┐
                                           DPᵢ     F′
                                                ┌──┴──┐
                                                F    PredP
                                                    ┌──┴──┐
                                                   Adverb ADJ
```

Sentences contained an animate subject and a verb that introduced a small clause (Contreras 1987; Jiménez-Fernández & Spyropoulos 2013). Within each small clause was a noun (NP1) modified by a prepositional phrase that contained a distractor noun (NP2); the head noun (NP1) served as the subject to a predicative adjective (or participle; we label this element ADJ).[8] For such a sentence to be grammatical, ADJ had to inflect for number and gender, agreeing with NP1.[9] In the design of our stimuli, we manipulated the number and gender of NP1, NP2, and ADJ, creating grammatical and ungrammatical sentences.

Consistent with previous work on agreement attraction that has identified an asymmetry between number features in determining agreement attraction (Bock & Miller 1991; Bock & Eberhard 1993; Vigliocco et al. 1995; 1996; Vigliocco & Nicol 1998; Bock et al. 2001; Hartsuiker et al. 2003; Alcocer & Phillips 2009; Bock et al. 2012; Acuña-Fariña et al. 2014; Jegerski 2016), we found that native speakers of Spanish exhibit clear effects of attraction for number: ungrammatical sentences with a singular NP1 but plural NP2 and ADJ were rated higher than other cases of ungrammaticality. Also consistent with the previous literature, we failed to find evidence of attraction for gender. Crucially, we found that within the potential attraction conditions (i.e., sentences with singular NP1, but plural NP2 and plural ADJ, as in (6a; b) above), conditions with feminine head nouns in which agreement attraction occurred in two features were rated significantly lower than conditions with agreement attraction in only one feature. The analogous difference for masculine head nouns was not significant. Figure 1 plots average ratings for these potential attraction conditions, thus representing the baseline.

We interpreted these results as evidence for a split model for number and gender in the native Spanish grammar. Native speakers treat attraction in two features separately from

---

[8] Here and below we refer to the head noun and distractor phrase pretheoretically as NPs; this is merely an expository step.

[9] See Fuchs et al. (2015) for the details of the norming study mentioned in Footnote 6 that was conducted to ensure that sentences were interpreted with ADJ modifying NP1, not NP2.
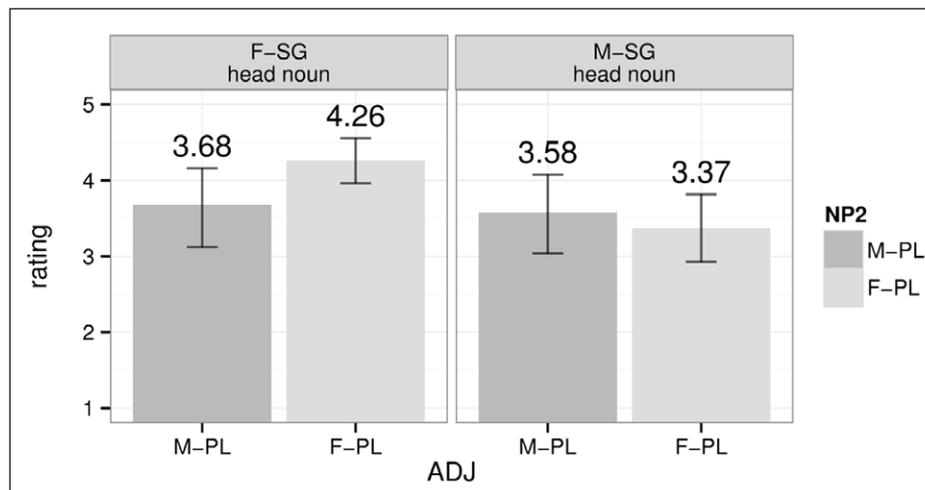
Figure 1: Average ratings for potential attraction conditions from the 50 native speakers of Spanish tested in Fuchs et al. (2015).

attraction in one. Because the two features participate in agreement separately, attraction in number does not precipitate attraction in gender.[10]

In addition to testing bundling vs. split models of number and gender, we set out to determine the content of each feature type. We used evidence from grammaticality effects—the ability to distinguish ungrammatical sentences from grammatical ones—in non-attraction sentences to determine feature content. Finding sensitivity to both singular and plural agreement errors, we concluded that number in the native baseline is a multi-valued feature (both singular and plural are specified feature values; cf. Sauerland 2003; Scontras 2013a; b). Finding sensitivity only to agreement errors involving feminine adjectives, we concluded that gender is single-valued (feminine is the specified feature value, while masculine serves as the absence of gender specification, consistent with the classic analysis proposed by Harris 1991; see also Kramer 2015).

Having mapped, as it were, the terrain of number and gender in the native Spanish baseline, we now shift the question to heritage speakers: does their knowledge of Spanish morphology match that of the native baseline? If it does not, what sorts of pressures likely led to the difference? Concretely, do heritage speakers show evidence of bundling or splitting number and gender? And, relatedly, how is each category specified: does number require that both singular and plural be specified? Does gender require that both masculine and feminine be specified? Or are single-valued oppositions sufficient? If heritage speakers are driven by pressures to increase analyticity, we would expect them to converge on a maximally-articulated feature representation where number and gender are split, with each feature value specified. Such a convergence would entail maintaining the structure observed in the native grammar and enhancing it with additional feature values.

If, however, representational economy influences the grammar of agreement in heritage speakers, we would expect them to converge on a lighter-weight, less-well articulated feature representation where number and gender are bundled together, and with the minimal number of feature values specified. Of course, there also remains the possibility that there is no difference between the native and heritage grammars of agreement, which we

---

[10] Using electrophysiological evidence, Barber & Carreiras (2005) arrive at a similar conclusion, namely that number and gender agreement are separate in monolingual (Peninsular) Spanish. Experimental work examining how number and gender are used in establishing pronoun antecedents has also provided evidence for a split model of number and gender in Italian (de Vincenzi 1999; de Vincenzi & di Domenico 1999).

assume as the null hypothesis (although this would run counter to the well-documented divergence of heritage language morphology surveyed above).

3 Testing heritage speakers

To investigate the content and structure of heritage agreement categories, we ran a direct replication of the auditory sentence-rating task from Fuchs et al. (2015), this time testing English-dominant heritage speakers of Spanish. Previous studies have looked at number or gender agreement in heritage speakers of Spanish, most commonly in agreement between two expressions that are adjacent (e.g., Montrul et al. 2008; 2014; Foote 2011; Montrul 2013); ours is the first to apply the agreement attraction paradigm in an attempt to diagnose the structure and content of these speakers' knowledge of their agreement morphology. Building on the results of previous studies, we find it imperative to increase the difficulty in comprehension by introducing linear (as well as structural) distance between the noun bearing agreement features and the agreeing expression. We also wanted to ensure that the target of agreement was the same for both gender and number (i.e., the predicative adjective), thus allowing for direct comparison across feature types (cf. Acuña-Fariña et al. 2014, who also compared two feature types).

If we find that the heritage speakers lose sensitivity to specific feature values, or that they bundle rather than split their number and gender features, then we have an instance of grammar divergence. More importantly, we can use any differences between the heritage and native grammars to help explain and thereby predict the gradual loss of agreement in heritage languages.

3.1 Methods

3.1.1 Participants

We recruited 160 participants with U.S. IP addresses via Amazon.com's Mechanical Turk crowd-sourcing service. Participants were compensated for their participation. The study was preceded by a demographics questionnaire. We identified as heritage speakers those participants who 1) grew up speaking Spanish, 2) now speak mostly English, and 3) had not lived in a Spanish-speaking country after the age of 8. Seventy-one participants matched the profile of English-dominant heritage speakers of Spanish; we present their results below.

Some high-proficiency heritage speakers of Spanish have been reported to perform on par with native speakers with respect to gender marking and agreement (Montrul et al. 2008; 2013; 2014; Alarcón 2011). Despite not including an independent measure of proficiency in an effort to keep the experiment short, the criteria used in our demographics questionnaire helped to exclude such high-proficiency speakers from the analysis, focusing instead on mid- to low-proficiency heritage speakers. We report the responses to these demographic questions in Appendix A: Demographic information. This information gives a fuller picture of our participants: most are young (average age is 31), speaking at least some Spanish at home, and never studied Spanish in college. There is a good deal of variability in the amount of grade-school Spanish these speakers were exposed to, and most report that they are native speakers of Spanish. The relatively high number of self-proclaimed native speakers points to the blurry line between heritage and native identification. These numbers also reinforce the observations in the literature that self-assessment by heritage speakers is sometimes inversely correlated with their proficiency (cf. Beaudrie & Ducar 2005; Thompson 2015, for Spanish heritage speakers; Titus 2012; Davidson & Lekic 2013, for Russian heritage speakers), or reflects the degree of their ethnic and cultural identification rather than proficiency (Kang & Kim 2012).

3.1.2 Materials

We used the same experimental stimuli from Fuchs et al. (2015). In designing both the previous and the present study, the goal was to construct stimuli that would allow us to examine how errors in agreement—specifically, errors in one vs. two agreement features—are perceived. As described in the previous section, we used the phenomenon of agreement attraction to effect these errors.

Stimuli took the form of the schema in (8), identical to (7) above; the schema features a small clause (SC) in which we expected to find cases of agreement attraction.

(8)   (Subject) Verb [$_{SC}$ **NP1** Prep NP2 Adv **ADJ**] …

We created sixteen items (for a full list of experimental stimuli, see Appendix B). For each item, grammatical and ungrammatical versions were derived by manipulating the number and gender features of NP1, NP2, and ADJ. Each of NP1, NP2, and ADJ could be either masculine or feminine, and either singular or plural, yielding four possible combinations of features per element. Grammatical configurations with masculine singular and feminine singular nouns appear in (9a) and (9b), respectively. Ungrammatical conditions were constructed by varying the features of NP1 and ADJ; potential attraction conditions were created by aligning the features of NP2 and ADJ.

(9)   a.   El   niño   considera   el   artículo   en el
           the.M.SG boy.M.SG consider.PRS.3SG   the.M.SG news.item.M.SG in the.M.SG
           periódico   terriblemente aburrido   pero le   gustaba
           newspaper.M.SG terribly   boring.M.SG but DAT.CL.3SG like.IPFV.3SG
           leer   los   libros   de Harry Potter.
           read.INF the.M.PL book.M.PL of Harry Potter
           'The boy considers the article in the paper to be terribly boring but he liked reading Harry Potter books.'

      b.   El   niño   considera   la   noticia   en la
           the.M.SG boy.M.SG consider.PRS.3SG   the.F.SG news.item.F.SG in the.F.SG
           revista   terriblemente aburrida   pero le   gustaba
           newspaper.F.SG terribly   boring.F.SG but DAT.CL.3SG like.IPFV.3SG
           leer   los   libros   de Harry Potter.
           read.INF the.M.PL book.M.PL of Harry Potter
           'The boy considers the article in the paper to be terribly boring but he liked reading Harry Potter books.'

Stimuli were recorded by a native speaker of Colombian Spanish. In order to avoid any prosodic cues to ungrammaticality, only grammatical sentences were recorded. All stimuli (both grammatical and ungrammatical) were created by splicing together grammatical [(Subj) Verb NP1 Prep NP2] onsets with grammatical [Adv ADJ …] offsets. We also created 16 fillers of similar complexity to our test items; the same speaker recorded both the test items and fillers. Given that heritage speakers often have problems with literacy, it was appropriate to rely on the auditory presentation of the stimuli.[11]

3.1.3 Design

Participants took the experiment online via Mechanical Turk using the web-based experiment platform ExperigenRT (Becker & Levine 2010; Pillot et al. 2012). Instructions were

---

[11] See Carreira (2007); Edstrom (2007); Valdés et al. (2008), for a discussion of literacy issues specifically with respect to heritage speakers of Spanish.

provided in written Spanish. Participants were presented with two versions at random of each of the 16 items (one version where the number of NP1 and NP2 matched and another where it did not), together with 16 fillers. Participants were instructed to indicate the acceptability of the sentences they heard on a scale from 1 (*completamente inaceptable* 'completely unacceptable') to 5 (*completamente aceptable* 'completely acceptable'). Participants completed a total of 48 trials: 32 test trials and 16 fillers presented in a random order.

3.2 Predictions

Table 1 summarizes the experimental predictions depending on whether pressures to increase analyticity or representational economy influence the heritage grammar of agreement; we repeat the results obtained by Fuchs et al. (2015) for the monolingual native baseline. Given that the monolingual baseline contains a multi-valued number feature, a single-valued gender feature, and a split feature representation where number and gender are projected and valued separately, the only way to increase the analyticity of this system is to move from a single- to a multi-valued gender feature. In other words, we should observe grammaticality effects for both masculine and feminine head noun conditions. However, if pressures from representational economy drive the heritage grammar's restructuring, we should observe fewer feature values and less structure: single-valued number and gender features that are bundled (i.e., projected and valued) together. Concretely, we should observe grammaticality effects for only one of the singular and plural head noun conditions, and for only one of the masculine and feminine head noun conditions. In potential attraction conditions where the features of the local noun (but not the head noun) match those of the agreeing adjective, we should observe that single-feature errors are rated as high or as low as multiple errors.

3.3 Results

To simplify presentation and ease interpretation, we follow Fuchs et al. (2015) and present the results in two parts. First, we investigate grammaticality and attraction effects for each of the four potential feature values (i.e., singular, plural, masculine, and feminine); analyses are grouped by the features of NP1 (i.e., the head noun), which ought to determine agreement. Second, we focus in on potential attraction conditions to test the predictions of bundled vs. split models of number and gender.

3.3.1 Head-noun analysis

Figure 2 plots average acceptability ratings grouped by the relevant feature of NP1.[12] Perhaps the most striking aspect of Figure 2 is the relatively high ratings across the board: on average, participants rated critical items 3.7 out of a possible 5 points on the acceptability scale. This may initially seem surprising, especially given that half of the conditions

Table 1: Experimental predictions presented against the monolingual baseline from Fuchs et al. (2015).

	monolingual baseline	increased analyticity	representational economy
number	multi-valued	multi-valued	single-valued
gender	single-valued	multi-valued	single-valued
bundled vs. split	split	split	bundled

---

[12] Responses were trimmed on the basis of response time to be within two standard deviations of the mean.
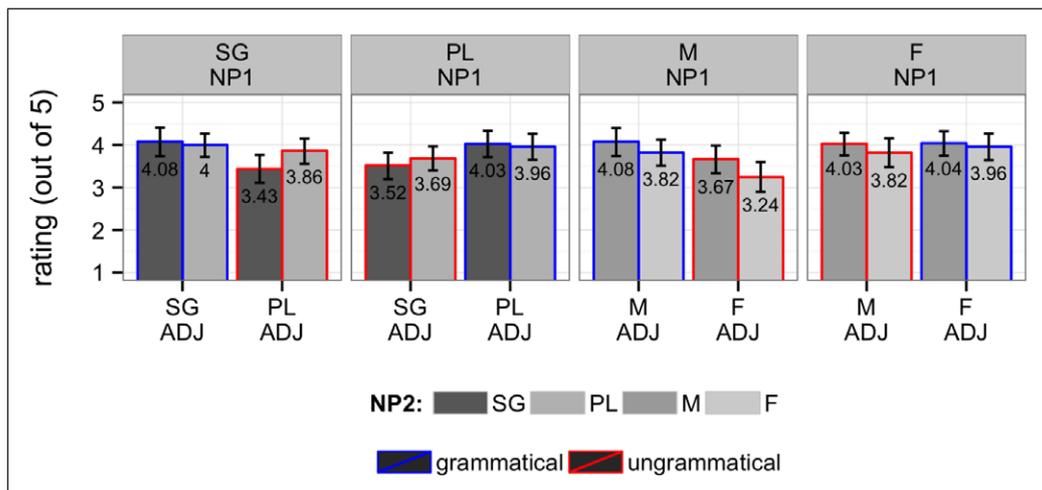
Figure 2: Average acceptability ratings grouped by NP1 features. Error bars represent bootstrapped 95% confidence intervals.

tested were ungrammatical. Moreover, previous research has demonstrated clear sensitivity to agreement violations in heritage speakers of Spanish (Foote 2011). However, recall that target ungrammaticality is the result of a mismatch in agreement features on NP1 and ADJ, and that these two constituents are contained within a much larger frame. The agreement error in ungrammatical conditions is just one small part of an otherwise lengthy and grammatically correct sentence. The fact that these agreement errors did not yield extremely low average ratings is therefore a product of the experimental design, which prioritized taxing participant memory by increasing distance between key sentential elements. This design yielded the same overall high ratings when the experiment was conducted with native speakers in Fuchs et al. (2015). The crucial metric in the previous and in the present study is therefore not the absolute value of an average rating, but rather how ratings compare to each other across different grammatical and ungrammatical conditions.

Considering first the most basic of these differences, heritage speakers gave grammatical conditions in the study an average rating of 3.95, and ungrammatical conditions an average rating of 3.66; this difference was found to be significant (Wilcoxon test, $p < 0.001$). This difference suggests that participants were attending to the experimental items and sensitive to differences in grammaticality. We turn next to further differences across various grammatical and ungrammatical conditions, grouping analyses by the relevant features of NP1.

Singular NP1 To avoid possible effects of gender mismatches, gender was held constant across NP1, NP2, and ADJ—all were either feminine or masculine. Keeping NP1 singular, there are four possible combinations for the number features of NP2 and ADJ, as in Figure 2. To investigate the effect of these feature combinations on the perceived acceptability of the sentences that contained them, we fit a linear mixed-effects model predicting acceptability ratings by NP2 (SG vs. PL) and ADJ (SG vs. PL), as well as their interaction. This model and the models that follow included random intercepts for participants and items. The model found a main effect of ADJ ($\beta = -0.74$, $t = -3.57$, $p < 0.05$): ungrammatical sentences in which ADJ appeared in the plural were rated lower than their grammatical counterparts. The model also found a marginally significant interaction between NP2 and ADJ ($\beta = 0.56$, $t = 1.93$, $p < 0.06$): when both NP2 and ADJ were plural, ungrammatical sentences were rated higher than the ungrammatical baseline. These results replicate the native baseline observed by Fuchs et al. (2015) and indicate that attraction from the plural is present in heritage speakers as well.

Plural NP1 As with the singular-NP1 analysis, we held gender constant. Keeping NP1 plural, there are four possible combinations for the number features of NP2 and ADJ, as in Figure 2. We fit a linear mixed-effects model predicting acceptability ratings by NP2 (SG vs. PL) and ADJ (SG vs. PL), as well as their interaction. The model did not find a significant effect of ADJ ($\beta = -0.27$, $t = -1.24$, $p < 0.22$): ungrammatical sentences in which ADJ appeared in the singular were not rated significantly lower than their grammatical counterparts. The lack of a significant difference between grammatical and ungrammatical conditions was a reflection of variation in participants' responses; in particular, participants were divided in their ratings of ungrammatical plural-NP1 conditions. These results deviate from the native baseline observed in Fuchs et al. (2015), where native speakers demonstrated a grammaticality effect for plural NP1. The model also did not find a significant interaction between NP2 and ADJ ($\beta = -0.22$, $t = -0.70$, $p < 0.49$): as with the native baseline, there is no evidence of attraction from the singular.

Masculine NP1 Turning now to gender, to avoid possible effects of number mismatches, number was held constant across NP1, NP2, and ADJ—all were either singular or plural. Keeping NP1 masculine, there are four possible combinations for the gender features of NP2 and ADJ, as shown in Figure 2. We fit a linear mixed-effects model predicting acceptability ratings by NP2 (M vs. F) and ADJ (M vs. F), as well as their interaction. The model found a significant effect of ADJ ($\beta = -0.46$, $t = -2.07$, $p < 0.05$): ungrammatical sentences in which ADJ appeared in the feminine were rated lower than their grammatical counterparts. The model did not find an interaction between NP2 and ADJ ($\beta = -0.02$, $t = -0.07$, $p < 0.95$): there is no evidence of attraction from the feminine. These results replicate the native baseline.

Feminine NP1 As with the masculine-NP1 analysis, we held number constant. Keeping NP1 feminine, there are four possible combinations for the gender features of NP2 and ADJ, as in Figure 2. We fit a linear mixed-effects model predicting acceptability ratings by NP2 (M vs. F) and ADJ (M vs. F), as well as their interaction. The model did not find a significant effect of ADJ ($\beta = -0.16$, $t = -0.73$, $p < 0.47$): ungrammatical sentences in which ADJ appeared in the masculine were rated no lower than their grammatical counterparts. The model also did not find an interaction between NP2 and ADJ ($\beta = 0.13$, $t = 0.44$, $p < 0.66$): there is no evidence of attraction from the masculine. These results replicate the native baseline.

3.3.2 Attraction condition analysis

Next, we focus in on potential attraction conditions where a singular NP1 occurs with a plural NP2 and ADJ. Figure 3 plots average responses to these conditions grouped by the gender of NP1.

We fit a linear mixed-effects model predicting ratings by NP1 (F-SG vs. M-SG), the number of agreement errors (1 vs. 2), and their interaction. The model found a main effect of NP1 ($\beta = -0.72$, $t = -2.40$, $p < 0.05$): attraction conditions with a feminine NP1 were rated significantly higher than attraction conditions with a masculine NP1. The model did not find an effect of the number of agreement errors ($\beta = -0.23$, $t = -0.77$, $p < 0.45$), or an interaction between the number of agreement errors and NP1 ($\beta = 0.42$, $t = 1.03$, $p < 0.31$).[13] Thus, for the feminine head nouns, conditions with agreement attraction in only one feature (F-SG F-PL F-PL) were rated as high as conditions with

---

[13] Given the lack of even a statistical trend (i.e., the difference between 4.07 and 4.08 in Figure 3), our failure to find a significant interaction with the number of agreement errors is unlikely to stem from a lack of statistical power.
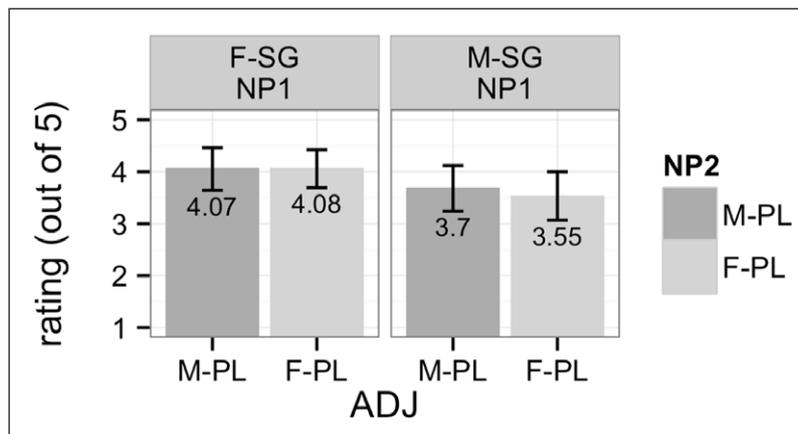
Figure 3: Average acceptability ratings for potential attraction conditions grouped by NP1 gender. Error bars represent bootstrapped 95% confidence intervals.

attraction in two features (F-SG M-PL M-PL); for masculine head nouns, conditions with agreement attraction in two features (M-SG F-PL F-PL) were rated as low as conditions with agreement attraction in only one feature (M-SG M-PL M-PL).

3.4 Discussion

We have extended our agreement-attraction paradigm based on auditory sentence-rating beyond the native baseline to English-dominant heritage speakers of Spanish. The results indicate both similarities and, crucially, differences between the two populations. First we consider the implications of the head-noun analyses for the content of the feature categories (i.e., number and gender) they inform. Then we turn to the interaction between the two categories, interpreting the results of the attraction analysis in light of the predictions from bundling vs. split theories of number and gender.

3.4.1 Number in heritage Spanish

Like the native baseline, heritage speakers showed a clear grammaticality effect for conditions in which the head noun was singular, as well as an effect of attraction from plural local nouns. However, *unlike* the native baseline, the grammaticality effect for plural head nouns was only a numerical trend, not significant. That is, participants did not consistently distinguish between grammatical sentences with a plural head noun and a plural adjective, and ungrammatical sentences with a plural head noun but a singular adjective. In other words, some participants perceived the agreement between a plural head noun and a singular adjective as grammatical.

If the representation of number in heritage Spanish were shaped by transfer from the dominant language's grammar, we would not expect this representation to be affected: native English number has the same properties as number in Spanish (Fuchs et al. 2015), and so transfer of properties from English to the heritage Spanish grammar should result in the preservation of the native number system. However, our results suggest that the content of the number category in the heritage grammar is moving away from the multi-valued native baseline, where both singular and plural are specified. Instead, the feature space is pruned in such a way that singular emerges as a kind of default, seemingly compatible with targets that are either specified (i.e., plural) or unspecified.

To explain such a restructuring of the number category in the heritage grammar, we hypothesize that a single-valued (privative) opposition, where only one of the members is specified for a particular feature, is somehow less taxing than the multi-valued opposition

where all the members have a predefined feature specification. While a multi-valued representation might seem preferable because it allows for a one-to-one mapping between features and morphology, a single-valued opposition decreases the memory load by allowing the speaker to only pay attention to and store in memory the specified value, in a sense omitting the unspecified value. This move in principle cuts the speaker's workload in half: the content of the category where only one feature is specified leads to a more economical representation.

3.4.2 Gender in heritage Spanish

With gender, heritage speakers behave on par with the native Spanish baseline: they demonstrate a clear grammaticality effect for masculine but not feminine head nouns; masculine adjectives may agree with feminine head nouns, signaling that masculine serves as the default gender specification—the absence of feminine. The baseline single-valued gender opposition in native Spanish appears to be robust enough to withstand pressures from general attrition and transfer from English—the heritage speakers' dominant language, which lacks grammatical gender—so that it persists in the heritage grammar. This suggests that a single-valued feature opposition does enjoy privileged status in the heritage grammar, further supporting our proposal concerning the single-valued opposition observed for heritage number.

3.4.3 Bundling vs. splitting in heritage Spanish

Turning now to the relationship *between* features, we found that participants rated potential attraction conditions with errors in one feature equally high (or low) as conditions with errors in two features. Thus, errors (or attraction) in number are perceived as equally (un)grammatical as errors in both number and gender, indicating that number and gender in heritage Spanish are bundled—that is, projected and valued—together. As such, in an agreement attraction context, when there is attraction in one feature (say, number), the other feature (say, gender) also participates in attraction, resulting in the lack of difference between the number of errors on perceived acceptability. This behavior stands as a departure from the native baseline, where we found that errors in both number and gender were perceived as less acceptable than errors in just number, where presumably attraction masks the offense. To summarize: we have found evidence that heritage speakers of Spanish bundle number and gender features in their grammar, a departure from the native baseline of feature splitting.

It bears noting that there may remain an obstacle to the conclusion that English-dominant heritage speakers of Spanish bundle their number and gender features: perhaps the lack of difference between attraction conditions with one error (i.e., number only) vs. two errors (i.e., number and gender) is not the result of any difference in the mental representation of these features, but rather of the lack of sensitivity to inflectional morphology that is often associated with heritage language. In other words, perhaps the observed insensitivity to the number of agreement errors signals not that number carries gender along for the ride while it gets valued in the heritage grammar, but rather that our heritage participants did not access gender as they processed the sentences presented to them. Were this true, our results would say nothing about the mental representation of the relationship between the two agreement features in question: the equal ratings given by the heritage speakers to conditions with agreement attraction in one vs. two features would be the result of the already-well-documented neglect of inflectional morphology.

We would like to offer a counterargument to this interpretation of the data. If heritage speakers were so insensitive to gender morphology that they had insufficient awareness of

gender errors, we would expect that they would treat attraction conditions with feminine head nouns on a par with conditions with masculine head nouns. However, this is not what we found: the feminine-head-noun attraction conditions are rated significantly higher than the masculine-head-noun attraction conditions. In fact, we already found evidence of gender sensitivity in the head-noun analyses, where we observed a clear grammaticality effect for masculine head nouns. This pair of findings—contrasts between feminine- and masculine- head-noun attraction conditions and grammaticality effects for masculine head nouns—indicates that heritage speakers are aware of and sensitive to gender. Indeed, Cuza & Pérez-Tattam (2016) find early knowledge of gender distinctions in child heritage learners of Spanish. If we take this evidence seriously, then we can maintain the conclusion that heritage speakers have reanalyzed the feature system of Spanish so that it levels the hierarchical distinction between number and gender. In other words, what native speakers treat as separate categories (i.e., number and gender), heritage speakers handle as but one, thus opting for the bundling of these categories. The featural organizations of number and gender in native and heritage Spanish are therefore ultimately different.

Given that our results come from testing elements separated by linear and structural distance, it seems reasonable to hypothesize that the observed difficulty in matching the features of distant constituents may start as a memory problem, yet another casualty of working memory demands on issues of dependency distance (Gibson 1998; Grodner & Gibson 2005). Heritage speakers have limited resources for integrating linguistic information, and these resources are taxed to the limit when information units are not contiguous. However, even though this difficulty may start as a performance issue (i.e., not domain-specific), it has deep consequences for the restructuring of the heritage grammar. In other words, the root cause of restructuring might exist external to language, but the consequences in the language domain are remarkably specific. We discuss these consequences in detail in the following section.

4 General discussion

A comparison of the results of the present study with those reported in Fuchs et al. (2015) points to grammatical divergence between native and heritage Spanish. First, concerning the content of number and gender features, English-dominant heritage speakers of Spanish appear to be losing sensitivity to the singular, treating it as the unspecified member of the opposition. The result is a single-valued featural representation of number in heritage Spanish where the native baseline has one that is multi-valued. Second, whereas the native baseline demonstrated differential sensitivity to single- vs. multi-error attraction conditions, heritage speakers lack a similar sensitivity and instead treat the two types of errors on a par. Recall that for heritage speakers in cases of number attraction, gender also participates; this indicates that the two features are projected and valued together, thus supporting a bundling model of their representation, as in (1) above. All told, with respect to gender and number features of Spanish, the native baseline and the heritage grammar show genuine divergence. Returning to our predictions from Table 1 in Section 3.2, our results support the hypothesis that pressures from representational economy drive the observed restructuring in the heritage grammar.

In what follows, we would like to offer three additional considerations. First, we will add a note of caution to the interpretation of our dependent measure. Then we will consider in more detail the pressures that likely led to the observed divergence between the heritage grammar and the native baseline. Finally, we will discuss the implications of our findings for the loss of morphological richness more generally.

4.1 A question of dependent measures

Although the performance of heritage speakers in the present study cannot be straightforwardly attributed to a lack of sensitivity to gender morphology, other external factors may affect the interpretation of our findings. A question arises—one that is common to experimental work—as to whether the methodology employed here accurately reflects the participants' knowledge of their language. Performance need not be an accurate representation of a speaker's knowledge, and so it may be with heritage speakers: their performance on a sentence-rating task need not give us full understanding of their competence in Spanish—other measures may bring to light additional evidence concerning their grammatical knowledge.

Evidence of the potential divide between grammaticality judgments and grammatical knowledge comes from Tokowicz & MacWhinney (2005). Picking up a lead from earlier work, Tokowicz & MacWhinney investigated discrepancies in the results of grammaticality judgments and event-related potentials (ERPs) gathered from L2 and native speakers of Spanish. The two groups were asked to judge the grammaticality of Spanish sentences that had one of three types of errors: 1) tense marking, which is similar in Spanish and English, 2) determiner-number agreement, which is different in Spanish and English, and 3) gender agreement, which exists in Spanish but not in English. ERP measures were recorded throughout the task. By comparing judgments to ERPs, the authors aimed to determine whether certain kinds of tasks reflected explicit versus implicit knowledge of a language.

While none of the L2 Spanish speakers tested by Tokowicz & MacWhinney was able to perform above chance at correctly identifying any of the three types of errors, the ERP results demonstrated more native-like performance. The L2 speakers displayed a P600 effect indicative of syntactic anomaly with tense-marking and gender-agreement error types; there was no observable grammaticality effect for number agreement. The authors interpreted these results as evidence that introspective judgments (similar to those we collected in our study) measure explicit knowledge of a language, while electrophysiological responses more directly evidence implicit, grammatical knowledge. Regardless of the specific interpretation, their results indicate that different dependent measures potentially differ in their sensitivity to grammatical knowledge. Given the potential limitations of acceptability judgments, future work on heritage populations (indeed, any population) ought to use corroborating evidence from different methodologies in the diagnosis of linguistic knowledge.[14] An obvious solution to this conundrum would be to compare methodologies in an explicit way and to combine data from production and comprehension as well.

Still, we do not believe that these considerations devoid the significance of our findings. We lack any evidence demonstrating that electrophysiological measures differ from behavioral ones in the present domain for heritage speakers. Moreover, the simple fact remains: heritage speakers pattern differently from the native baseline when it comes to the perceived acceptability of agreement-attraction sentences. In the following sections, we argue that the precise nature of this pattern points to a principled path of grammar divergence.

---

[14] See Montrul (2016: Chapter 6); Orfitelli & Polinsky (2017), for a similar line of argumentation, namely that grammaticality judgment tasks potentially obscure grammatical knowledge in near-native speakers. Fortunately, we avoid the main pitfall that they identify with grammaticality judgment tasks by moving from binary judgments (i.e., acceptable vs. unacceptable) to graded judgments in the current study.

4.2 Developmental trajectories of heritage language

Having documented principled differences between heritage and native speakers of Spanish with respect to the structuring of gender and number, the task now is to understand why and how the observed differences arise. First, let us take stock of what we have observed: 1) a move toward single-valued feature oppositions in both number and gender (as compared to the multi-valued number opposition in the native baseline), and 2) the bundling of number and gender features (as opposed to the split-feature model of the native baseline). Let us now we revisit the predictions from Table 1 and the potential pressures that might have led to the restructuring that we observe in the heritage grammar: representational economy vs. increased analyticity.

If heritage speakers were prioritizing analyticity in their morphosyntax, we might expect them to converge on a grammar that privileges one-to-one correspondences between surface forms and their underlying features (cf. the observed loss of ambiguity in word order, Isurin & Ivanova-Sullivan 2008; Ivanova-Sullivan 2014; in syntactic representations, Polinsky 2011; 2016b; and in scope interpretations, Scontras et al. 2017). With respect to feature content, we would expect the heritage grammar to be at least as well-articulated (i.e., to have at least as many feature values specified) as the native baseline. Indeed, the grammar should have more feature values specified (i.e., multi-valued oppositions in both number and gender). We make a similar prediction with respect to the feature representation: the heritage grammar should have a representation with at least as much articulated structure as the native baseline. In both cases, greater articulation of feature content and structure increases analyticity. However, neither prediction is borne out in our results; we find exactly the opposite: a loss of feature content and a loss of structure.

Heritage speakers appear instead to be prioritizing representational economy, presumably to ease the load on their working memory as they carry out the costly task of using a less dominant language. This restructuring converges on a grammar with more parsimonious linguistic representations: fewer features and less structure overall. The result is the grammar we observe here, where agreement categories are smaller and bundled together, stored and retrieved in memory as one unit. One might worry that the observed primacy of representational economy stands at odds with the previously-reported results of increased analyticity: how do we reconcile this case with the many others where analyticity appears to rule the day? Here it bears noting that in the other domains of grammar that we have reported on, increased analyticity and representational economy most likely deliver supporting—or orthogonal—pressures. Take the simplification of gapped-relativization structures documented by Polinsky (2011) for English-dominant heritage speakers of Russian. The resulting heritage grammar has fewer relativization options than the native baseline; the single viable option that remains is the relativization of subjects with a gap. If any of the relativization structures has a lighter-weight structure, it would be the subject relative clause, which minimizes syntactic dependencies. At worst, representational economy has no bearing on the outcome of relativization restructuring; at best, it delivers the same result as the pressures from increased analyticity: subject relative clauses. So we would have no way of teasing apart the relative strength of the two pressures. However, in the domain of morphosyntactic agreement discussed here, our operationalizations of representational economy vs. increased analyticity make diverging predictions, and we have seen that our results readily align with the predictions of representational economy.

It took a case where the two pressures push in opposite directions to observe the dominance of representational economy in the heritage grammar of agreement. In the next section, we discuss the implications of these findings for the loss of morphological richness more broadly construed.

4.3 Implications for loss of morphological richness

The differences in the underlying syntax of agreement categories in heritage as compared to native Spanish suggest a clear trajectory for the development of the impoverished agreement systems characteristic of heritage grammars and, more broadly, of grammars under contact.

In the native Spanish grammar, number and gender are projected independently along the nominal spine, with NumP either dominating GenP, according to the schema in (2) above (which is repeated below in (10a)), or with gender specified on the nominal stem, as in (10b) (Ritter 1993; Alexiadou 2004; Kramer 2015; 2016).[15] Thus, native speakers have a fully articulated structure for their nominal projections, and because each agreement category is independent, each feature also has a clear feature specification.

(10)   a. $[_{DP} \ldots [_{NumP\{xy\}} \ldots [_{GenP\{xy\}} \ldots [_{nP} \ldots$ *gender as separate XP*
       b. $[_{DP} \ldots [_{NumP\{xy\}} \ldots [_{nP\{+Gender\}} \ldots$ *gender specified on stems*

This is a rich system where structure building and parsing have to scan a fairly large structure to build a DP. Morphological richness follows, and the inflectional nature of Spanish entails that various pieces of morphology may be spelling out several consecutive categories.

Contrast this picture with what we find in heritage speakers, who combine number and gender in a single projection. Whereas some work has suggested that gender features in a language cannot appear solely on NumP (Kramer 2015: Chapter 8), our findings suggest that this is indeed the case in the grammar of heritage Spanish. Assuming that the structure in (10a) is widely available as a default (cf. Harley & Ritter 2002), this bundling may follow from the fusion of the Gen head with the Num head. Departing from the structure in (10a), the new feature bundle is projected on NumP, and GenP is left empty (and may subsequently be deleted from the structure altogether).[16] As a result, heritage speakers lose feature specification on one of the nodes in the DP:

(11) $[_{DP} \ldots [_{Num\text{-}GenP\{xy\}} \ldots [_{GenP\{EMPTY\}} \ldots [_{nP} \ldots$

Assuming the alternative representation of number and gender in the baseline grammar as in (10b), bundling occurs because gender collapses into NumP and is no longer associated with the nominal stem:

(12) $[_{DP} \ldots [_{Num\text{-}GenP\{xy\}} \ldots [_{nP} \ldots$

Arguments as to which structure, (10a) or (10b), is the underlying one are quite complex and are beyond the scope of this paper. The results of our study are equally compatible with both approaches (i.e., gender as separate XP and gender as specified on stems), and we are not in a position to adjudicate between the two. It is possible that both structures are made available in Universal Grammar, and considerations for one or the other are language-specific. Crucially, for our purposes the end result is the same: the emergence of a bundled node with number and gender represented together. In sum, while the baseline grammar splits number and gender, the heritage grammar represents gender

---

[15] Alternatively, gender may be specified on the root, not on n. We will not consider this option here; for our purposes, it is not meaningfully different from (10b).

[16] The morphological implementation may vary by language and the specific approach taken—options include head raising (cf. our discussion of example (3) above) and PF lowering (Embick & Noyer 2001).

only on NumP, leading to a general reduction of the featural system. We now consider the consequences of this reduction.

When number and gender are separate, their featural content is maximally transparent, but when they are bundled together their featural content becomes far more opaque. Although representational economy favors a feature bundle, as discussed above, this structure requires more processing effort to disentangle the individual features. A speaker no longer recognizes number and gender, but rather some hazy amalgamation of the two. This feature opacity may lead to interpretive instability.

At this stage, the process can go one of two ways. Given interpretive instability, the feature bundle might be reinterpreted as one feature only—for instance, number might be preserved, to the exclusion of gender, which will be lost. This outcome is advantageous as it results in a feature whose semantic content is once again clear. Alternatively, the feature bundle might lose feature specification altogether, resulting in an empty feature projection, as in (13) below. The outcome is a more general decline in morphological richness, eventually leading to the loss of agreement. This latter outcome is also a hallmark of heritage languages (Benmamoun et al. 2013b); now we have a better understanding of how this loss transpires.

(13) $[_{DP} \ldots [_{NumP\{EMPTY\}} \ldots [_{nP} \ldots$

The trajectory outlined here demonstrates that loss of agreement in heritage languages is not accidental, but rather follows from changes in feature representation and specification. When considering the syntactic structure of the native baseline, different pressures suggested different outcomes for the heritage grammar. As in Table 1 (Section 3.2), if changes in the featural system in heritage Spanish were driven by the pressure for increased analyticity often found in other domains of heritage grammars, then we would expect number and gender to remain split in the heritage grammar and even become more analytic by having a more fine-grained (i.e., multi-valued) feature specification. However, our results argue against this general tendency toward increased analyticity: the heritage featural system is instead guided by representational economy, as it bundles together features that are split in the baseline grammar. This tension between the desire for increased analyticity and decreased opacity on one hand, and a more trim syntactic structure on the other hand may be the driving force behind the loss of certain features in heritage grammars, with gender in heritage Spanish serving as an example of the outcome of this tension.

As described above, the gradual impoverishment of morphological richness is driven by systematic pressures that can be predicted on the basis of syntactic structures. One concrete prediction made by this system is that person features should be more resilient to loss of agreement than agreement in number and gender in the verbal domain. Recall from the discussion of the phi-feature hierarchy that person is known to be syntactically separate from number and gender features and is probed for prior to the probing for the other phi-features. As person features do not have a parallel structural relationship to either number or gender, the conditions for bundling are not as easily met. As a result, person features should not be as susceptible to bundling and loss of morphological richness.

5 Conclusions

We have used differences in linguistic behavior—differential response patterns to acceptability judgments for potential attraction conditions—to diagnose differences in linguistic knowledge of gender and number in adjectival agreement at a distance. Previous studies have explored agreement in heritage Spanish (Montrul et al. 2008; 2013; 2014; Alarcón

2011; Montrul 2016), but ours is the first to add linear and structural distance between the agreeing goal and probe while extending the agreement attraction paradigm to heritage speakers. This innovation serves to provide more information about the heritage grammar of agreement while simultaneously suggesting a clear path of restructuring from the monolingual baseline to the heritage grammar.

Our main findings are twofold. First, we show that whereas native Spanish has a multi-valued opposition in the number category (singular and plural are equally specified), heritage speakers are moving toward a single-valued opposition, losing sensitivity to the singular. Second, we find that whereas native speakers of Spanish split their number and gender features, heritage speakers bundle them, projecting and valuing the two features together. The restructuring that leads to this difference in knowledge likely proceeds under systematic pressure from representational economy: heritage speakers shrink their feature inventory and prune their structure so that it includes fewer nodes. A possible competing force is the desire for increased analyticity in the heritage grammar, but pressures from increased analyticity made predictions not supported by the experimental results; this in turn suggests that increased analyticity is not the critical goal of the heritage system's restructuring in this domain. Crucially, the restructuring of gender and number in heritage Spanish is not a random process, but rather proceeds according to the structural constraints of the grammar in which it operates. The bundling of number and gender is possible because these feature categories are projected close to each other on the nominal spine; no structural rearrangement is necessary for these features to merge.

Returning to the observations that initially motivated the current study, we have found that within Spanish, speakers with varying levels of proficiency possess diverging but related knowledge of the Spanish grammar. Performing the costly task of speaking in a less-dominant language, heritage speakers show evidence of a grammar of agreement that has succumbed to pressures from representational economy. But should we conclude from this finding that those heritage speakers are speaking a *different* language from the native baseline: that they know a language other than Spanish? Our claim is that we should not. Both groups are speaking Spanish. However, there are different types of Spanish speakers (indeed, speakers of any language) based on the amount of input, maintenance, and interference from other languages. The resulting structuring of grammar in these speakers may differ, as we showed in this paper, but it is a coherent grammar, one that can be accommodated and predicted under the existing models of linguistic variation. Furthermore, the emergence of principled differences of the type documented here underscores the need for comprehensive models of *intra*-linguistic variation.

This result speaks to the advancing desire to include heritage speakers' knowledge under the rubric of native speakerhood; such a desire is particularly pronounced in sociolinguistic work, sometimes compelling researchers to reject the term "heritage language" altogether (e.g., Otheguy 2016: 309–311). Although we do not wish to abandon the conception of heritage language, which has served us well and has stimulated a great deal of sophisticated research, we believe that the work presented here gives novel support to the idea that heritage speakers can be included in the intra-linguistic variation within native grammars.

Finally, having documented pressures from representational economy—in addition to the well-established pressures to increase analyticity—shaping the heritage grammar, a new question looms large: is it possible to predict which language domains may be likely to restructure to deliver more parsimonious representations, and which domains may be likely to restructure to increase analyticity? We do not yet have answers to this question, but the results presented in this work compel us to pose it.

Additional Files

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

- **Appendix A.** Demographic information. DOI: https://doi.org/10.5334/gjgl.164.s1
- **Appendix B.** Experimental items. DOI: https://doi.org/10.5334/gjgl.164.s1

Abbreviations

1 = first person, 2 = second person, 3 = third person, ADJ = adjective, ADV = adverb, CL = clitic, COND = conditional, DAT = dative, F = feminine, IPFV = imperfective, INF = infinitive, M = masculine, NEG = negation/negative, PL = plural, PRS = present, SG = singular

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

References

Acuña-Fariña, J. Carlos, Enrique Meseguer & Manuel Carreiras. 2014. Gender and number agreement in comprehension in Spanish. *Lingua* 143. 108–128. DOI: https://doi.org/10.1016/j.lingua.2014.01.013

Alarcón, Irma. 2011. Spanish grammatical gender under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition* 3. 332–350. DOI: https://doi.org/10.1017/S1366728910000222

Alcocer, Pedro & Colin Phillips. 2009. A cross-language reversal in illusory agreement licensing. *Poster presented at the 22nd annual CUNY conference on human sentence processing.* University of California, Davis.

Alexiadou, Artemis. 2004. Inflection class, gender and DP-internal structure. In Gereon Müller, Lutz Gunkel & Gisela Zifonum (eds.), *Explorations in nominal inflection,* 21–50. Berlin: Mouton de Gruyter.

Antón-Méndez, Inés, Janet L. Nicol & Merrill F. Garrett. 2002. The relation between gender and number agreement processing. *Syntax* 5. 1–25. DOI: https://doi.org/10.1111/1467-9612.00045

Armoskaite, Solveiga. 2011. *The destiny of roots in Blackfoot and Lithuanian.* Vancouver, BC: University of British Columbia dissertation.

Atkinson, Emily A. 2015. Gender features on *n* and the root: An account of gender in French. In Jason Smith & Tabea Ihsane (eds.), *Romance linguistics 2012: Selected papers from the 42nd linguistic symposium on Romance languages (LSRL),* 229–244. Amsterdam: Benjamins.

Au, Terry Kit-Fong, Janet Oh, Leah Knightly, Sun-Ah Jun & Laura Romo. 2008. Salvaging a childhood language. *Journal of Memory and Language* 58(4). 998–1011. DOI: https://doi.org/10.1016/j.jml.2007.11.001

Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

Barber, Horacio & Manuel Carreiras. 2005. Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience* 17(1). 137–153. DOI: https://doi.org/10.1162/0898929052880101

Barlow, Jessica A. 2014. Age of acquisition and allophony in Spanish-English bilinguals. *Frontiers in Psychology* 5(APR). 1–14. DOI: https://doi.org/10.3389/fpsyg.2014.00288

Beaudrie, Sara & Cynthia Ducar. 2005. Beginning level university heritage programs: Creating a space for all heritage language learners. *Heritage Language Journal* 3(1). 1–26.

Becker, Michal & 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 speaker: Theoretical and methodological challenges Reply to peer commentaries. *Theoretical Linguistics* 39(3–4). 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(3–4). 129–181. DOI: https://doi.org/doi.org/10.1515/tl-2013-0009

Bock, Kathryn & Carol A. Miller. 1991. Broken agreement. *Cognitive Psychology* 23(1). 45–93. DOI: https://doi.org/10.1016/0010-0285(91)90003-7

Bock, Kathryn & Kathleen M. Eberhard. 1993. Meaning, sound, and syntax in English number agreement. *Language and Cognitive Processes* 8. 57–99.

Bock, Kathryn, Kathleen M. Eberhard, J. Cooper Cutting, Antje S. Meyer & Herbert Schriefers. 2001. Some attractions of verb agreement. *Cognitive Psychology* 43. 83–128.

Bock, Kathryn, Manuel Carreiras & Enrique Meseguer. 2012. Number meaning and number agreement in English and Spanish. *Journal of Memory and Language* 66(1). 17–37. DOI: https://doi.org/10.1016/j.jml.2011.07.008

Bolonyai, Agnes. 2016. (in)vulnerable agreement in incomplete bilingual l1 learners. *International Journal of Bilingualism* 11(1). 3–23.

Bull, William. 1965. *Spanish for teachers: Applied linguistics*. New York, NY: Ronald Press.

Carminati, Maria Nella. 2005. Processing reflexes of the Feature Hierarchy (person > number > gender) and implications for linguistic theory. *Lingua* 115(3). 259–285. DOI: https://doi.org/10.1016/j.lingua.2003.10.006

Carreira, Maria. 2007. Spanish-for-native-speaker matters: Narrowing the Latino achievement gap through Spanish language instruction. *Heritage Language Journal* 5(1). 147–171.

Carstens, Vicki. 2000. Concord in minimalist theory. *Linguistic Inquiry* 31. 319–355.

Chang, Charles B. 2016. Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition* 19. 791–809.

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

Clahson, Harald & Claudia Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27. 3–42.

Contreras, Heles. 1987. Small clauses in Spanish and English. *Natural Language & Linguistic Theory* 5. 225–243.

Croitor, Blanca & Ion Giurgea. 2009. On the so-called Romanian "neuter". *Bucharest Working Papers in Linguistics* 11(2). 21–39.

Cuza, Alejandro & Rocío Pérez-Tattam. 2016. Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition* 19(1). 50–68.

Davidson, Don E. & Maria Lekic. 2013. The heritage and non-heritage learner in the overseas immersion context: Comparing learning outcomes and target-language utilization in the Russian flagship. *Heritage Language Journal* 10(2). 88–114.

de Vincenzi, Marica. 1999. Differences between the morphology of gender and number: Evidence from establishing coreferences. *Journal of Psycholinguistic Research* 28. 537–553.

de Vincenzi, Marica & Elisa di Domenico. 1999. A distinction among ϕ-features: The role of gender and number in the retrieval of pronoun antecedents. *Rivista di Linguistica* 11. 41–74.

Edstrom, Anne. 2007. The mixing of non-native, heritage, and native speakers in upper level Spanish courses: A sampling of student opinion. *Hispania* 90. 755–768.

Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595.

Fathi, Radwa & Jean Lowenstamm. 2016. The gender assignment pattern of French nouns. *Morphology* 26. 477–509.

Ferreira, Fernanda, Karl G. D. Bailey & Vittoria Ferraro. 2002. Good enough representations in language comprehension. *Current Directions in Psychological Science* 11. 11–15.

Foley, William. 1991. *The Yimas language of New Guinea*. Stanford: Stanford University Press.

Foote, Rebecca. 2011. Integrated knowledge of agreement in early and late English–Spanish bilinguals. *Applied Psycholinguistics* 32. 187–220.

Fuchs, Zuzanna, Maria Polinsky & Gregory Scontras. 2015. The differential representation of number and gender in Spanish. *The Linguistic Review* 32(4). 703–737. DOI: https://doi.org/10.1515/tlr-2015-0008

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.

Giurgea, Ion. 2008. Gender on definite pronouns. *Bucharest Working Papers in Linguistics* 10(1). 97–121.

Godson, Linda. 2003. *Phonetics of language attrition: Vowel production and articulatory setting in the speech of Western Armenian heritage speakers*. San Diego, CA: University of California, San Diego dissertation.

Grodner, Daniel & Edward Gibson. 2005. Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science* 29(2). 261–290.

Guitart, Jorge M. 1997. Variability, mulitlectalism, and the organization of phonology in Caribbean Spanish dialects. In Fernando Martínez-Gil & Alonso Morales-Front (eds.), *Issues in the phonology and morphology of the major Iberian languages*, 515–536. Washington DC: Georgetown University Press.

Håkansson, Gisela. 1995. Syntax and morphology in language attrition: A study of five bilingual expatriate Swedes. *International Journal of Applied Linguistics* 5(2). 153–169.

Harley, Heidi. 1994. Hug a tree: Deriving the morphosyntactic feature hierarchy. *Massachusetts Institute of Technology Working Papers in Linguistics* 21.

Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3). 482–526.

Harris, James W. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22(1). 27–62.

Hartsuiker, Robert J., Herbert J. Schriefers, Kathryn Bock & Gerdien M. Kikstra. 2003. Morphophonological influences on the construction of subject-verb agreement. *Memory and Cognition* 31. 1316–1326.

Hewitt, George B. 1979. *Abkhaz*. Amsterdam: North Holland.

Isurin, Ludmila & Tanya Ivanova-Sullivan. 2008. Lost in between: The case of Russian heritage speakers. *Heritage Language Journal* 6. 72–104.

Ivanova-Sullivan, Tanya. 2014. *Theoretical and experimental aspects of syntax-discourse interface in heritage grammars*. Leiden: Brill.

Jegerski, Jill. 2016. Number attraction effects in near-native Spanish sentence comprehension. *Studies in Second Language Acquisition* 38. 5–33. DOI: https://doi.org/10.1017/S027226311400059X

Jiménez-Fernández, Ángel & Vassilios Spyropoulos. 2013. Feature inheritance, vP phrases and the information structure of small clauses. *Studia Linguistica* 67. 185–224.

Kang, Hyun-Sook & In-sop Kim. 2012. Perceived and actual competence and ethnic identity in heritage language learning: A case of Korean-American college students. *International Journal of Bilingual Education and Bilingualism* 15(3). 279–294.

Keating, Gregory D., Jill Jegerski & Bill Vanpatten. 2016. Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition* 19(1). 36–49. DOI: https://doi.org/10.1017/S1366728914000418

Kramer, Ruth. 2009. *Definite markers, phi-features, and agreement: A morphosyntactic investigation of the Amharic DP*. Santa Cruz, CA: University of California, Santa Cruz dissertation.

Kramer, Ruth. 2015. *The morphosyntax of gender*. Oxford: Oxford University Press.

Kramer, Ruth. 2016. The location of gender features in the syntax. *Language and Linguistics Compass* 10. 661–677.

Kupisch, Tanja. 2012. Specific and generic subjects in the Italian of German-Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition* 15. 736–756.

Laleko, Oksana & Maria Polinsky. 2013. Marking topic or marking case? A comparative investigation of heritage Japanese and heritage Korean. *Heritage Language Journal* 10(2). 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

Lipski, John. 1984. On the weakening of /s/ in Latin American Spanish. *Zeitschrift für Dialektologie und Linguistik* 51. 31–43.

Lipski, John. 2011. Socio-phonological variation in Latin American Spanish. In Manuel Díaz-Campos (ed.), *The handbook of Hispanic sociolinguistics*, 72–97. London: Wiley-Blackwell.

Matushansky, Ora. 2013. Gender confusion. In Lisa L.-S. Cheng & Norbert Corver (eds.), *Diagnosing syntax*, 271–294. Oxford: Oxford University Press.

Meisel, Jürgen. 2007. The weaker language in early child bilingualism: Acquiring a first language like a second language? *Applied Psycholinguistics* 28. 495–514.

Montrul, Silvina. 2008. *Incomplete acquisition in bilingualism. Re-examining the age factor*. Amsterdam: John Benjamins.

Montrul, Silvina. 2013. How native are heritage speakers? A look at gender agreement in Spanish. *Heritage Language Journal* 10. 15–39.

Montrul, Silvina. 2016. *The acquisition of heritage languages*. Cambridge: Cambridge University Press.

Montrul, Silvina, Israel de la Fuente, Justin Davidson & Rebecca Foote. 2013. The role of experience in the acquisition and production of diminutives and gender in Spanish:

Evidence from L2 learners and heritage speakers. *Second Language Research* 29(1). 87–118.

Montrul, Silvina, Justin Davidson, Israel de la Fuente & Rebecca Foote. 2014. Early language experience facilitates gender agreement processing in Spanish heritage speakers. *Bilingualism: Language and Cognition* 17(1). 118–138.

Montrul, Silvina, Rebecca Foote & Silvia Perpiñán. 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(3). 503–553. DOI: https://doi.org/10.1111/j.1467-9922.2008.00449.x

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

Otheguy, Ricardon. 2016. The linguistic competence of second-generation bilinguals: A critique of "incomplete acquisition". In Christina Tortora, Marcel den Dikken, Ignacio L. Montoya & Teresa O'Neill (eds.), *Romance linguistics 2013: Selected papers from the 43rd linguistic symposium on Romance languages (LSRL), New York, 17–19 April, 2013*, 301–320. Amsterdam: John Benjamins.

Phillips, Colin. 2013. Some arguments and non-arguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes* 28(1–2). 156–187. DOI: https://doi.org/10.1080/01690965.2010.530960

Picallo, M. Carme. 1991. Nominals and nominalization in Catalan. *Probus* 3. 279–316.

Picallo, M. Carme. 2007. On gender and number. Ms., Universitat Autonoma de Barcelona.

Picallo, M. Carme. 2008. Gender and number in Romance. *Lingue e Linguaggio* 7. 47–66.

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. 1995. Cross-linguistic parallels in language loss. *Southwest Journal of Linguistics* 14(1–2). 87–123.

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

Polinsky, Maria. 2008a. Heritage language narratives. In Donna Brinton, Olga Kagan & Susan Bauckus (eds.), *Heritage language education: A new field emerging,* 149–164. New York, NY: Routledge.

Polinsky, Maria. 2008b. Relative clauses in heritage Russian: Fossilization or divergent grammar. In *Formal approaches to Slavic linguistics #16: The Stony Brook meeting 2007*, 5–21. Ann Arbor, MI: Michigan Slavic Publications.

Polinsky, Maria. 2011. Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition* 33. 305–328.

Polinsky, Maria. 2016a. Bilingual children and adult heritage speakers: The range of comparison. *International Journal of Bilingualism*. DOI: https://doi.org/10.1177/1367006916656048

Polinsky, Maria. 2016b. Structure vs. use in heritage language. *Linguistics Vanguard* 2(1). DOI: https://doi.org/10.1515/lingvan-2015-0036

Ritter, Elizabeth. 1993. Where's gender? *Linguistic Inquiry* 24. 795–803.

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

Sanford, Anthony J. & Patrick Sturt. 2002. Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences* 6. 382–386.

Sauerland, Uli. 2003. A new semantics for number. In Robert B. Young & Yuping Zhou (eds.), *Proceedings of SALT 13*, 258–275. Ithaca, NY: CLC Publications.

Scontras, Gregory. 2013a. Accounting for counting: A unified semantics for measure terms and classifiers. In Todd Snider (ed.), *Proceedings of SALT 23*, 549–569.

Scontras, Gregory. 2013b. A unified semantics for number marking, numerals, and nominal structure. In Emmanuel Chemla, Vincent Homer & Grégoire Winterstein (eds.), *Proceedings of Sinn und Bedeutung 17*, 545–562.

Scontras, Gregory, Maria Polinsky, C.-Y. Edwin Tsai & Kenneth Mai. 2017. Cross-linguistic scope ambiguity: When two systems meet. *Glossa: A journal of general linguistics* 2(1). 36. DOI: https://doi.org/10.5334/gjgl.198

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

Sherkina-Lieber, Marina, Ana T. Pérez-Leroux & Alana Johns. 2011. Grammar without speech production: The case of Labrador Inuttitut heritage receptive bilinguals. *Bilingualism: Language and Cognition* 14(3). 301–317.

Song, Minsun, William O'Grady, Sookeun Cho & Miseon Lee. 1997. The learning and teaching of Korean in community schools. In Y.-H. Kim (ed.), *Korean language in America 2*, 111–127. Washington DC: American Association of Teachers of Korean.

Sorace, Antonella. 2005. Syntactic optionality at interfaces. In Leonie Cornips & Karen Corrigan (eds.), *Syntax and variation: Reconciling the biological and the social*, 46–111. Amsterdam: John Benjamins.

Steriopolo, Olga & Martina Wiltschko. 2010. Distributed GENDER hypothesis. In Gerhild Zybatow, Philip Dudchuk, Serge Minor & Ekaterina Pshehotskaya (eds.), *Formal studies in Slavic linguistics: Proceedings of the formal description of Slavic languages 7.5*, 155–172. New York, NY: Peter Lang.

Thompson, Gregory. 2015. Understanding the heritage language student: Proficiency and placement. *Journal of Hispanic Higher Education* 14(1). 82–96.

Titus, Julia. 2012. Creating an online assessment test for heritage learners of Russian. *Heritage Language Journal* 9(2). 85–107.

Tokowicz, Natasha & Brian MacWhinney. 2005. Implicit and explicit measures of sensitivity to violations in second language grammar: An eventrelated potential investigation. *Studies in Second Language Acquisition* 27(2). 173–204. DOI: https://doi.org/10.1017/S0272263105050102

Valdés, Guadalupe. 2000. Introduction. In *Spanish for native speakers* I. 1–32. New York, NY: Harcourt College.

Valdés, Guadalupe, Joshua Fishman, Rebecca Chávez & William Pérez. 2008. Maintaining Spanish in the United States: Steps toward the effective practice of heritage language reacquisition/development. *Hispania* 91. 4–24.

Vigliocco, Gabriella, Brian Butterworth & Carlo Semenza. 1995. Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language* 34. 186–215.

Vigliocco, Gabriella, Brian Butterworth & Merrill F. Garrett. 1996. Subjectverb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition* 61(3). 261–298. DOI: https://doi.org/10.1016/S0010-0277(96)00713-5

Vigliocco, Gabriella & Janet L. Nicol. 1998. Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition* 68. B13–B29.

OPEN ACCESS

OPEN ACCESS