Open Library of Humanities

Failed gender agreement in L1 English L2 Spanish: Syntactic or lexical problem?

Alena Kirova, Youngstown State University, United States, akirova@ysu.edu
José Camacho, University of Illinois Chicago, United States, jcamach@uic.edu

A recent proposal attributes morphosyntactic issues in L2 to *lexical* factors (Grüter et al. 2012; Hopp 2013). According to this lexical account, issues with gender agreement are caused by gender *assignment* issues – a failure to assign a word to a target-like class. We elaborate on this idea by exploring three potential cues to gender assignment: 1) semantic gender relating to sex (e.g. 'girl' vs. 'boy') 2) morphophonological cues, and 3) morphosyntactic agreement cues. Semantic and morphophonological cues may facilitate gender agreement only for a subset of nouns, whereas agreement cues can do so for all nouns, including *opaque* gender nouns that do not have semantic gender.

Seventeen low proficiency and sixteen high proficiency L1 English L2 Spanish speakers and eighteen native Spanish controls judged the grammaticality of 60 experimental sentences. We compared participants' gender agreement accuracy and reaction times (RTs) on experimental items with and without semantic gender, and with and without transparent gender morphemes. Semantic gender did not serve as a cue for gender assignment/agreement; instead, it slowed down RTs in high proficiency and control participants. Morphophonological cues significantly increased accuracy and decreased RTs in all groups. Finally, agreement cues did not seem to help low proficiency learners, since their accuracy on opaque nouns was barely above chance. By contrast, high proficiency learners exhibited native-like accuracy on opaque nouns. These findings support the lexical accounts of L2 gender agreement difficulties, adding more data to the growing body of research in this field.

1 Introduction

Grammatical gender is an inherent *lexical* feature on noun roots that triggers *syntactic* operations within the nominal domain (Carstens 2000; 2010). Unlike with first languages (L1), second language (L2) acquisition of grammatical gender has been shown to be notoriously difficult (Carroll 1989; Hawkins & Franceschina 2004; Franceschina 2005; Hawkins 2009, inter alia), with even fairly advanced learners making gender errors (Bruhn de Garavito & White 2002; Franceschina 2005; Montrul et al. 2008; Alarcón 2011; Grüter et al. 2012). In this paper, we explore lexical factors that may contribute to L2 learners' poor performance on grammatical gender by testing low and high proficiency L1 English L2 Spanish speakers on a timed grammaticality judgement task.

On the one hand, some representational approaches, such as the Failed Functional Features Hypothesis (FFFH) (Hawkins & Chan 1997) and the Representational Deficit Hypothesis (Tsimpli & Dimitrakopoulou 2007; Hawkins 2009), have ascribed L2 gender acquisition difficulties to impairment at the deep syntactic level and proposed that functional features such as grammatical gender cannot be acquired after a certain critical period. On the other hand, other representational approaches, like the Missing Surface Inflection Hypothesis (MSIH) (Haznedar & Schwartz 1997; Prévost & White 2000), have attributed these difficulties to production errors: specifically, the inability to *produce* the right gender morpheme in real-time. Yet another influential approach is based on Lardiere (2008; 2009) Feature Reassembly Hypothesis (FRH), which suggests that variability in the L2 performance on gender is not an issue of availability of features in the L2 grammars but rather an issue of remapping feature configurations from the L1 to different configurations in the L2 – a complex process that raises challenges that may account for non-target aspects of the L2.

Several processing accounts have also proposed that poor performance on syntactic structures in L2 does not stem from a lack of syntactic representations. For example, the Shallow Structure Hypothesis (Clahsen & Felser 2006) has regarded these difficulties as issues having to do with *underusing* syntactic representations online rather than with a lack of such representations. McDonald (2006) has shown that issues with syntactic structures on a grammaticality judgment task are positively correlated with individual differences in such non-grammatical cognitive processes as working memory capacity and decoding ability. Finally, Hopp (2014: 251) proposed that issues with syntactic processing may arise from the "overburdening of a native-like processing system" due to allocating particularly high amounts of cognitive resources to lexical processing.

In fact, as far as specifically grammatical gender is concerned, Grüter et al. (2012) and Hopp (2013; 2016) have put forward a hypothesis according to which gender agreement errors observed in L2 learners are a result of *lexical* factors and are not necessarily indicative of problems with syntactic processing or with underlying syntactic representation. In support of this hypothesis, a number of more recent psycholinguistic and neurolinguistic studies indicate that non-target like gender agreement in L2 results when a speaker assigns a word to the non-target-like gender

class (i.e. a lexical assignment divergence, cf. Sabourin & Stowe 2008; Alarcón 2011; Grüter et al. 2012; Hopp 2013; 2016).

That is, many errors that appear to be errors of gender agreement – a grammatical operation that distributes the noun gender (such as feminine or masculine) among different elements of the noun phrase (e.g. determiners, adjectives, quantifiers, etc.) (Carstens 2000; Pesetsky & Torrego 2004) – are in fact caused by (incorrect) gender assignment – assigning of a noun to a given gender class (e.g. Spanish *mesa* 'table' → feminine) (Corbett 1991). Since gender agreement is triggered by the inherent gender of the noun (*mesa* is inherently and invariably of feminine gender), it can only be fully native-like if all nouns of the gendered language have been assigned to the correct (L1-like) gender class in the speaker's mental lexicon.[1] Even if gender agreement is present in the L2 learner's interlanguage, but a specific noun has not been assigned to a gender class or has been assigned to a different gender by the L2 learner, it may lead to non-target-like gender agreement. For example, if the Spanish noun *mano* 'hand.F', which is feminine in L1, has been assigned to the masculine gender, it will lead to non-target-like gender agreement *el mano* '*DEF.M hand.M.' Such an error, however, is not a gender *agreement* mismatch but a gender *assignment* mismatch, and therefore a lexical, not a syntactic, issue (Grüter et al. 2012).[2]

The gender assignment process – in which a speaker assigns nouns to a specific gender class in the lexicon – may proceed on the basis of two types of cues found in the input: extralinguistic cues, such as semantic gender, and intralinguistic cues, such as morphophonological cues and agreement cues (Pérez-Pereira 1991; Spinner & Juffs 2008; Spinner & Thomas 2014). Semantic (or natural) gender is a feature of animate nouns that refers to the biological sex of the entity, e.g. 'girl' vs. 'boy'. In Spanish, most animate nouns that have a female referent in the real world belong to a feminine gender class, while most nouns with a male referent belong to a masculine gender class, and this semantic cue may facilitate assigning nouns to the two gender classes in the lexicon. By contrast, most inanimate nouns like *mesa* 'house.F' are assigned a purely grammatical arbitrary gender feature and do not have a matching counterpart with the opposite gender. We will refer to this purely grammatical feature as *arbitrary* gender.

Morphophonological (or form-based, having to do with the *form* of the word) cues are gender morphemes on nouns such as the *–a* morpheme on feminine gender nouns (*casa* 'house-F') and the *–o* morpheme on masculine gender nouns (*queso* 'cheese-M') in Spanish. Agreement (or structural/morphosyntactic) cues are gender morphemes on determiners (*el* 'DEF.M' vs. *la* 'DEF.F') and adjectives (*blanc-a* 'white-F' vs. *blanc-o* 'white-M'). These latter cues are instrumental

---

[1] This is not to say that L2 learners are incapable of performing gender agreement; rather, they cannot perform gender agreement in a fully *native-like* way, unless all L2 nouns are assigned to the target gender.

[2] Note that having assigned nouns to the correct gender class is a necessary, but not a sufficient condition for target-like gender agreement. Even if the learner has correctly *assigned* the word *mano* to a feminine gender, she may still make a gender *agreement* error.

for assigning gender to nouns that have opaque gender morphemes and lack semantic gender. For example, since the morpheme –*e* on a noun like *puente* 'bridge' is uninformative with respect to gender and the noun does not have semantic gender, the only way a beginning language learner (either L1 or L2) can assign the noun to a gender class is through being exposed to input containing an agreement cue, as in *el puente* 'DEF.M bridge.?' → *puente*.M.

While it has been shown that L1 learners use all of these cues to assign gender (Karmiloff-Smith 1979; Pérez-Pereira 1991; Müller 1994), it is not clear whether L2 learners make use of all or only some of these cues, or, crucially, how their use or non-use of these cues affects their grammatical gender acquisition.

In this paper, we test the effect of three potential gender assignment cues on gender agreement in low and high proficiency L1 English L2 Spanish speakers in pursuit of the following three main goals. First, we explore the effect of semantic gender by comparing L2 gender agreement accuracy and reaction times on experimental items that have semantic gender (e.g. *chica* 'girl') and those that do not (e.g. *casa* 'house'). Second, we study the effect of morphophonological form (i.e. transparency or opacity of the gender morpheme) by comparing L2 gender agreement accuracy and reaction times on experimental items that have transparent gender morphemes such as –*o* for masculine (*ques-o* 'cheese-M') and –*a* for feminine (*cas-a* 'house-F') and on nouns that have opaque gender morphemes such as –*z* (*lápiz* 'pencil.M') and –*e* (*fuente* 'fountain.F'). These manipulations allow us to investigate the effect of lexical factors on gender agreement, and thus to add data to the existing literature on lexical causes of gender agreement difficulties (assigning gender to nouns) in L2 learners.

Third, we compare low and high proficiency L1 English L2 Spanish speakers as well as a control group of native speakers of Spanish specifically with regard to their performance on *opaque* gender nouns. In doing so, not only do we show how a lexical characteristic of Spanish gender (i.e. opacity of gender morphemes) can affect L2 performance on gender agreement, but also use this lexical characteristic as a tool to test acquirability of gender agreement in L2. English lacks grammatical gender; thus, according to deficit approaches to L2 morphosyntax, L2 learners whose L1 is English should not be able to fully acquire gender in the L2. If L2 learners are indeed incapable of using agreement cues due to a permanent maturational constraint, our L1 English L2 Spanish participants should not be able to assign gender to nouns with opaque gender morphemes, because agreement cues are the only cues available to assign gender to opaque nouns. Gender agreement errors should hence be present on opaque gender nouns in the two proficiency groups (although expectedly to a lesser extent in the high proficiency one), and, importantly, even the high proficiency speakers should not behave like controls in terms of accuracy and reaction times.

Just as earlier studies have shown the different ways in which factors other than syntactic deficit may affect L2 morphosyntax (e.g. mapping features to forms, limited cognitive resources, etc.), we explore the way lexical characteristics of noun gender affect gender agreement. The paper

is organized as follows: first, we define the basic terms used in this paper, such as grammatical gender, gender assignment, and gender agreement; then, we describe the three types of gender assignment cues and review existing relevant literature; next, we posit our research questions, present our results and discussion of those results; and, finally, we lay out our conclusions.

2 Grammatical gender: Assignment and agreement in Spanish

Grammatical gender on Spanish nouns is a two-way classificatory system with no clear semantic content in the case of nouns with inanimate referent: *casa* 'house' is feminine (F), but *apartamento* 'apartment' is masculine (M), *silla* 'chair' is feminine, but *sillón* 'armchair' is masculine. As a consequence, every noun's lexical entry must encode grammatical gender that must be learned as part of the meaning of the word, possibly as part of a lemma or the lexical-syntactic representation of the word (Caramazza 1997; Levelt et al. 1999). This process (and the result) of incorporating language-specific genders (M and F) of nouns into the lexicon by language learners is referred to as gender *assignment*.

By contrast, gender *agreement*, according to linguistic theories (Carstens 2000; Pesetsky & Torrego 2004), represents a grammatical operation that distributes the noun's inherent (lexical) gender among different elements of the noun phrase that do not have an inherent gender but have functional gender features (e.g. determiners, adjectives, quantifiers, etc.). According to the analysis in Pesetsky and Torrego (2004), agreement is driven by unvalued features on elements that need valuation during the syntactic derivation. More specifically, the same feature (e.g. grammatical gender) can be distributed among different syntactic elements (e.g. nouns, determiners, adjectives) at different syntactic locations, and some of the features (e.g. gender feature on the noun) are valued (i.e. have a value such as feminine or masculine), while others (e.g. gender feature on determiners and adjectives) are unvalued (i.e. do not have a value). In this case, the unvalued feature (F [ ]) will have to find the identical, but valued feature (Fval) to enter in an agreement relationship with and to obtain the value from it. The unvalued feature is said to *probe* the goal – the valued feature. Gender values of Spanish nouns are stored in the lexicon and are specified as masculine or feminine. Modifiers and determiners within the determiner phrase (DP) have gender features as well, but these are unvalued; hence, they search their c-command domain for an element (noun) with the same, but valued, feature. They probe the valued gender feature in order to get valuation and eventually to delete. Thus, the gender feature is lexically specified on a noun (e.g. *mesa* 'table-F', but not 'table-M'), and both determiners and adjectives vary systematically depending on the gender of noun:

(1)   a. El   sill-ón   negr-o
             DEF.M armchair-M black-M
             'The black armchair'

    b.   La   sill-a   negr-a
         DEF.F chair-F black-F
         'The black chair'

Since it is the noun's gender that is shared with the other elements of the noun phrase, gender agreement cannot be *target-like* unless the noun is assigned to a correct gender class. That is, determiners and adjectives cannot receive a gender from the noun that they modify if the noun has not already been assigned to a gender. Thus, assigning gender to the noun is a prerequisite for *target-like* gender agreement.

Given the principal differences between gender assignment and gender agreement coupled with their highly interrelated nature, finding a practical solution to distinguish between agreement and assignment errors is an important goal, and various studies have used different methodological approaches to it. For example, some studies have taken agreement between the noun and the determiner to indicate gender assignment, and agreement between the noun and the adjectives to indicate gender agreement (Dewaele & Véronique 2001; Kupisch et al. 2013). Under such an approach, a DP such as *el mariposa' * DEF.M butterfly-F' would be taken to indicate a gender assignment error, while a DP such as *la mariposa rojo '*DEF- F red-M butterfly-F' – a gender agreement error. Other authors (e.g. Bruhn de Garavito & White 2002; White et al. 2004), have considered both agreement between the noun and the determiner and the noun and the adjective as indicators of gender agreement. The latter approach is more consistent with the definition of gender agreement, which is a feature-checking or feature-valuation relationship between the inherent gender features on the noun, on the one hand, and functional gender features on determiners and adjectives, on the other hand (Carstens 2000; Pesetsky & Torrego 2004).

Alarcón (2004), Montrul (2008), and Grüter et al. (2012) have used an insightful way to distinguish between assignment and agreement errors: they classified errors with either the determiner or the adjective as an agreement error (2a), and errors with both the determiner and the adjective as an assignment error (2b). In addition, Grüter et al. (2012) classified errors with an incorrect determiner and correct adjective as potentially both an agreement and assignment error (2c). This is undoubtedly a more robust way of teasing apart assignment and agreement errors; however, one may argue that even a DP such as (2b) probabilistically could be an agreement error or an error that reflects a strategy of defaulting to masculine rather than an assignment error.

(2)   a.   Agreement error: la   mariposa rojo
                               DEF-F butterfly-F red-M
                               'red butterfly'

        b.   Assignment error: el   mariposa rojo
                               DEF.M butterfly-F red-M

    c.   Both: el   mariposa   roja
                            DEF.M butterfly-F red-M
                            'red butterfly'

For the purposes of this study, we are only looking at DPs containing determiners and nouns and, following Bruhn de Garavito & White (2002) and White et al. (2004), consider lack of agreement between these two elements to be a gender agreement error. What is of interest to us here is how gender agreement between the two elements is affected by the lexical properties of the noun, such as transparency vs. opacity and the presence or absence of semantic gender.

As was suggested in the Introduction, gender assignment can be incrementally acquired based on three types of cues. In the following section, we will discuss these in detail.

2.1 Semantic cues

Semantic (natural) gender in animate nouns like *chica* 'girl-F' or *chico* 'boy-M' might serve as a cue for L2 learners to assign gender to nouns. L2 research in this area is not conclusive – while some studies indicate that semantic gender facilitates L2 gender processing (Finnemann 1992; Fernández-García 1999; Alarcón 2009), other studies suggest that gender agreement on nouns with semantic gender are processed more slowly and less accurately than on nouns with no semantic gender (Bruhn de Garavito & White 2002; Sagarra & Herschenson 2011; Sagarra & Herschenson 2012).

One reason why semantic gender may serve as a cue is that it is aligned with the noun's grammatical gender – animate nouns referring to female entities also carry feminine grammatical gender, and those referring to male entities carry masculine grammatical gender. In addition, specifically in Spanish, animate nouns referring to female entities (e.g. *hermana* 'sister', *prima* 'female cousin', *sobrina* 'niece', etc.) very often have an *–a* morpheme, which most frequently coincides with the feminine gender, while those referring to male entities (e.g. *hermano* 'brother', *primo* 'male cousin', *sobrino* 'nephew', etc.) very often have an *–o* morpheme, which coincides with masculine gender (cf. Harris 1991). The association of gender morphemes on these nouns with clear semantic gender features may make the morphemes salient and thus easier to acquire. Unlike inanimate nouns like *mesa* 'table-F' that don't have a semantic correlate for gender, animate nouns do have at least an indirect one: the grammatical gender of most of them corresponds to their (natural) semantic gender.

On the other hand, the reason why semantic gender might *not* serve as a cue or could even adversely affect gender agreement is that activation of an animate noun with semantic gender leads to activation of two lexical items – those denoting female and male entities (e.g. activating *hermano* 'brother' will also activate *hermana* 'sister', activating *chico* 'boy' will activate *chica* 'girl', etc.), and thus may cause semantic interference (Levelt et al. 1999; Indefrey & Levelt 2004), which the processor has to resolve by eventually choosing one item. This is not the case for the

majority of inanimate nouns; for example, activating a word like *queso* 'cheese' does not involve activating a word like **quesa*; hence, there is less competition and, therefore, less processing cost for activating nouns with semantic gender than nouns with arbitrary gender.

Furthermore, marking semantic gender on a noun helps distinguish between real world entities belonging to different sexes, whereas gender agreement has a purely grammatical function – that of facilitating and accelerating prediction of an upcoming noun in the speech stream (Bates et al. 1995; Akhutina et al. 1999). An L2 learner may process or produce the *–a* ending on the noun *chica* 'girl' to establish that this word refers to a female entity; however, this does not necessarily mean that she will also extend this semantic piece of information to the linguistic property of gender *agreement* such as *la chica* 'DEF.F girl-F'. Anecdotally, L2 learners very often produce utterances such as **el chica* 'DEF.M girl-F'. That is, the necessity of processing/ expressing a semantic feature such as semantic gender can be satisfied with mere marking on the noun as in *chica* vs. *chico* (on the noun itself – word-*internally*); gender agreement, on the other hand, can only be satisfied if gender is successfully marked on all the elements of the DP (beyond the noun – word-*externally*) in a target-like way.

2.2 Morphophonological cues

Another cue to assign gender to nouns is the morphophonological cue – gender-informative (transparent) morphemes on nouns such as the masculine *–o* ending on the noun *queso* 'cheese' (phonological markers) or the *–dad* morpheme on the noun *felicidad* 'happiness' (morphological markers) (Kupisch et al. 2013). Although Spanish, like Italian, Russian and Hebrew, has a more transparent and clear gender system than languages such as German and Norwegian (Bordag et al. 2006), it does not exhibit strict gender deduction rules but rather shows probabilistic tendencies or patterns – for example, there is no single morpheme on nouns that would invariably mark the masculine or feminine class (cf. Harris 1991). Nevertheless, some morphemes can very reliably predict noun gender, and so are called *transparent*, while others cannot, and so are called *opaque*.

The two transparent morphemes that serve as the most reliable cues for deducing Spanish gender are the *–o* morpheme for masculine and the *–a* morpheme for feminine, because 99.89% of Spanish nouns that end in *–o* are masculine and 96.6% of nouns that end in *–a* are feminine (Teschner & Russell 1984). It must be noted, however, that some common words ending in *-a/-o* are not feminine and masculine respectively, as in (3).

(3)   a.   Feminine N in *–o*: *moto* 'motorbike', *mano* 'hand'
        b.   Masculine N in *–a*: *programa* 'program', *día* 'day'

In addition, only 62% of masculine Spanish nouns end in *–o*, while only 55.9% of feminine Spanish nouns end in *–a* in the Davies (2002) corpus (Martin et al. 2017), which means that there

are many nouns that end in morphemes other than these two. Thus, while *–a* and *–o* are strong predictors of grammatical gender in Spanish, there are many nouns whose gender they cannot predict.

At the same time, there are other transparent gender morphemes that may help assign gender to nouns and, therefore, serve as cues to gender assignment. For example, the *–dad* morpheme marks feminine gender in 97.57% of cases (Teschner & Russell 1984). Examples of other transparent morphemes include *–ez* and *–ción* for feminine and *–or*, *–ón*, and *–je* for masculine, as in (4).

(4)   a.   Feminine N: *fals-e-dad* 'falsity', *igual-dad 'equality', estrech-ez* 'narrowness', *pequeñ-ez* 'smallness', *dona-ción* 'donation', *ac-ción* 'action', etc.

        b.   Masculine N: *am-or* 'love', *fact-or* 'factor', *mont-ón* 'pile', *tabl-ón* 'plank', *dopa-je* 'doping', *monta-je* 'set up', etc.

On the other hand, *opaque* gender morphemes cannot be used to deduce the gender of nouns. For example, endings such as *–z, –is,* and *–l* can mark both feminine, as in (5), (6), (9), and (11) and masculine, as in (7), (8), (10), (12), and (13) and, therefore, are uninformative with respect to gender.

(5)   La   paz
        DEF.F peace.F
        'the peace'

(6)   La   nariz
        DEF.F nose.F
        'The nose'

(7)   el   lápiz
        DEF.M pencil.M
        'the pencil'

(8)   el   disfraz
        DEF.M disguise.M
        'the disguise'

(9)   la   dosis
        DEF.F dose.F
        'the dose'

(10)   el   análisis
        DEF.M analysis.M
        'the analysis'

(11) la   piel
      DEF.F skin.F
      'The skin'

(12) el   cartel
      DEF.M sign.M
      'The sign'

(13) el   pastel
      DEF.M cake.M
      'The cake'

Moreover, a large set of nouns end in *–e,* which is either ambiguous or in case of animate nouns can signal either masculine or feminine, as in (14).

(14) a. Masculine in *–e*: *monte* 'mount', *puente* 'bridge', etc.
      b. Feminine in *–e*: *fuente* 'fountain', *muerte* 'death', etc.
      c. Either in *–e*: *el/la duende* 'the(M/F) elf', *el/la presidente* 'the(M/F) president', *el/la representante* 'the(M/F) representative'

Hence, while transparent gender morphemes can serve as (fairly) reliable cues to gender class of the noun, opaque ones cannot. Thus, grammatical gender assignment can be facilitated by the former type of morphemes, but not by the latter.

The linguistic differences between transparent and opaque nouns lead to important differences between the way they are processed in the brain of the native speaker. For example, an ERP (event-related potential) study by Caffarra et al. (2014) has shown that native speakers use two different routes – a lexical route and a form-based route – to access the gender of opaque nouns and the gender of transparent nouns respectively. This finding supports Gollan & Frost's (2001) dual route hypothesis according to which the gender of nouns can be retrieved via a lexical route as an abstract feature stored in the lexicon and via a form-based route using the sub-lexical units (gender morphemes) that typically mark a specific gender class (e.g. *–o* typically marks masculine and *–a* typically marks feminine). According to the study results, when formal cues (i.e. having to do with the *form* of the word) are available (such as transparent gender morphemes on nouns), native speakers quickly utilize them to access the gender of the noun, but when they are not available (as in opaque nouns), they use the lexical route to do so. Other authors suggest that formal cues are redundant for native speakers, and that they mainly use the more reliable lexical route to access gender (Vigliocco & Franck 1999).

The results reported in Caffarra et al. (2014) are similar to findings of neuroimaging research, which has convincingly demonstrated that transparent and opaque nouns are

processed in different brain areas (Miceli et al. 2002; Indefrey & Levelt 2004; Padovani et al. 2005; Heim et al. 2006; Hammer et al. 2007; Heim 2008; Quiñones et al. 2018), and that processing opaque nouns happens at a deeper level and involves more cognitive effort. For example, a recent fMRI study by Quiñones et al. (2018) showed that processing of transparent nouns, which have phonological/orthographic cues, involves both brain areas responsible for morphological decomposition (such as the occipito-temporal and parietal brain areas) as well as areas responsible for lexical processing (such as the fronto-temporal and parietal areas). On the other hand, processing of opaque nouns, which do not have formal cues, requires retrieving gender from the lexicon and involves brain areas such as the posterior part of the MTG/STG, the pars triangularis within the IFG, and the hippocampus, as well as the parietal areas.

The ERP and neuroimaging findings outlined above are valid for native speakers, but they may not apply to L2 learners. For example, L2 learners have been shown to be less accurate on opaque gender nouns than on transparent gender nouns (Cain et al. 1987; Finnemann 1992; Fernández-García 1999; Schlig 2003; White et al. 2004; Franceschina 2005; Alarcón 2006, for L2 Spanish; Taraban & Roark 1996, for L1 English L2 French; Oliphant 1998, for L1 English L2 Italian; Taraban & Kempe 1999, for L1 English L2 Russian; Bordag 2004; Bordag et al. 2006, for L1 Czech L2 German; Kupisch et al. 2013, for L1 German L2 French and L1 French L2 German; Spinner & Thomas 2014, for L1 English L2 Swahili), which indicates that they may not be able to use the lexical route to access gender of opaque nouns, over-relying instead on the form-based route. This may be because L2 learners may not have assigned all nouns to a target-like gender class and thus their lexical entry for a given noun may lack an abstract gender feature.

Importantly, the finding that L2 learners are more accurate on transparent nouns is compatible with the idea that at least some gender agreement errors are in fact gender assignment errors, since it shows that L2 learners are capable of performing gender agreement on nouns when they have a means of assigning gender. In the next section, we will offer an explanation as to why L2 learners have difficulty building a lexical representation for nouns with opaque gender morphemes.

2.3 Agreement cues

Since grammatical gender cannot be assigned based on opaque gender morphemes, it can only be assigned to opaque gender nouns based on agreement cues in the input – the gender of the elements that modify the noun (e.g. the determiner or adjective). For example, one can assign the noun *puente* 'bridge', whose gender morpheme *–e* is ambiguous and thus uninformative with respect to the noun's gender, to the masculine gender class only by being exposed to input where the noun co-occurs with a masculine determiner (or another modifying element), as in *el puente* 'DEF.M bridge.?' → *puente*.M.

Therefore, failure to exploit agreement cues will often lead to failed gender agreement on opaque nouns, but not on transparent nouns – precisely the pattern L2 learners have exhibited

in L2 gender research. Importantly, since such failed gender agreement on opaque nouns would be caused by a lack of ability to establish a forward dependency relation between the gendered determiner and the noun – which is essentially a definition of gender agreement (Hopp 2016) – it would be indicative of compromised or immature gender agreement.

This is compatible with representational deficit accounts (Hawkins & Chan 1997; Tsimpli & Dimitrakopoulou 2007; Hawkins & Casillas 2008; Hawkins 2009) and specifically with Hawkins and Franceschina (2004), who argue that young L1 learners and L2 learners perform gender agreement *superficially*, that is, based on overt gender morphology of the noun, as in (15), whereas adult native speakers perform it *maturely*, that is, based on gender of the modifiers, as in (16).

(15)   *Superficial* agreement: gender morphemes on the noun → gender on modifiers

(16)   *Mature* agreement: gender on modifiers → noun gender

Crucially, according to the authors, while for L1 learners this superficial gender agreement represents a developmental stage, it represents an end-state grammar for L2 learners. In other words, while children eventually switch from the superficial, form-based, agreement to mature, syntax-based, agreement and thus reach full mastery of gender, L2 learners never do, unless their L1 has gender.

This hypothesis has been partially supported by L1 gender research. For example, L1 gender acquisition studies show that children prioritize morphological rules over syntactic rules at earlier ages, but shift to syntactic rules at older ages, roughly after the age of 5 (Karmiloff-Smith 1979, for French; Pérez-Pereira 1991, for Spanish; Müller 1994, for German and French). Moreover, eye-tracking research has shown that native speakers consistently use gender on articles or adjectives to accelerate and facilitate prediction of an upcoming noun during discourse processing (Dahan et al. 2000, for French; Sekerina et al. 2006, for Russian; Lew-Williams & Fernald 2007, for Spanish; Grüter et al. 2012, for Spanish; Hopp 2013, for German). Even toddlers as young as 2 or 3 years old can make use of agreement cues (gender-marked determiners) to predict upcoming nouns (Johnson 2005, for Dutch; Lew-Williams & Fernald 2007, for Spanish; van Heugten & Johnson 2011, for Dutch; Smolík 2014, for Czech).

Conversely, while the hypothesis advanced in Hawkins and Franceschina (2004) is compatible with the findings of L2 research on transparent vs. opaque nouns outlined above, findings of eye-tracking research with L2 learners are not conclusive. For example, some studies found that moderately proficient L2 learners (Lew-Williams & Fernald 2010) and even highly proficient L2 learners (Grüter et al. 2012) fail to exploit gender on determiners and adjectives to predict upcoming nouns and, thus, would be unlikely to use such cues to assign gender to nouns. At the same time, other L2 studies (Hopp 2013; Dussias et al. 2013) have shown that L2 learners *are* capable of using agreement cues to construct predictive agreement relations between the

determiner and the noun. Hopp (2013) conducted an eye-tracking study with 20 advanced to near-native L1 English L2 German speakers. He found that a subgroup of L2 participants who had successfully assigned nouns to a target-like gender class was able to use gender on articles to predict the noun. Crucially, since German's gender system is not phonologically transparent, this effect could not have been the result of a superficial gender agreement (based on use of morphophonological cues), but rather truly reflected ability to use gender information on elements other than the noun – agreement cues. This finding runs contrary to representational deficit theories. Similarly, in an eye-tracking study Dussias et al. (2013) showed that L2 learners do not lack an ability to use agreement cues for predictive processing, but rather that this ability is modulated by proficiency and L1-L2 similarity. More specifically, while low proficiency Spanish speakers were not able to use gender on Spanish articles to facilitate the processing of upcoming nouns, highly proficient L1 English L2 Spanish speakers as well as Italian-Spanish bilinguals, whose L1 and L2 have similar gender systems, were indeed able to do so.

To summarize this section, L1 learners use all available cues in the input – morphophonological (form-based) cues, agreement (structural/morphosyntactic) cues and semantic cues – to assign gender and process gender agreement. Having used a combination of these cues while acquiring their L1 allows them to build two routes to access the gender of nouns – lexical and form-based routes – and thus to quickly retrieve the gender of both transparent and opaque gender nouns. They also routinely utilize agreement cues (gender on modifiers) to predict the upcoming noun in speech stream, as visual world paradigm (eye-tracking) studies demonstrate. Since L2 learners' ability to use agreement cues seems to be limited compared to NS, and given that gender can only be assigned to opaque nouns based on these cues, it is not surprising that L2 learners perform agreement less accurately on opaque nouns. This limited ability to use agreement cues represents a morphosyntactic issue, not a lexical one, since it has to do with gender agreement between the noun and the determiner, but manipulating a lexical characteristic of the noun – opacity of the noun gender – allows us to find out whether this limited morphosyntactic ability is a transient stage or a permanent state in L2 grammar.

3 The study

Some linguistic approaches to L2 acquisition, which study how language features are represented in one's linguistic system, argue that abstract linguistic representations cease to be available after a certain critical period (deficit approaches such as the Failed Functional Features Hypothesis and the Representational Deficit Hypothesis), while other linguistic approaches suggest that they are available through adulthood (non-deficit approaches such as the Feature Reassembly Hypothesis and the Missing Surface Inflection Hypothesis). L2 difficulties with gender agreement have been taken as an indication of an absence of abstract syntactic representations and thus as evidence in favor of representational deficit approaches. In this study, following Grüter et al. (2012) and

Hopp (2013), and given that gender is part of the noun's *lexical* representation, we suggest that some of those difficulties may be caused by lexical factors such as opacity vs. transparency of the noun gender morpheme. We also suggest that they may be caused by an overreliance on morphophonological cues (transparent gender morphemes) that results in a reduced tendency to attend to agreement cues (gender morphemes on determiners); however, following non-deficit approaches such as FRH, we argue that this tendency represents a temporary rather than permanent state of L2 grammars.

3.1 Hypotheses and research questions

We pursue three goals in this study. First, we test the effect of semantic gender (if any) by comparing L2 learners' performance on gender agreement on inanimate nouns and on animate nouns with semantic gender. Second, we compare gender agreement on opaque and transparent nouns in two language proficiency groups to replicate the results of the existing literature according to which L2 learners perform gender agreement less accurately on opaque nouns, thereby exhibiting overreliance on form-based cues and also a lack of abstract lexical representation of gender. Lastly, we compare gender agreement with opaque gender nouns in low and high proficiency learner groups to find out whether the L2 gender agreement limitation manifested in poor performance on opaque nouns represents just a developmental stage or a permanent state of L2 grammars. That is, if high proficiency learners, unlike low proficiency learners, are capable of performing gender agreement on opaque gender nouns, we can conclude that they have been able to utilize agreement cues and, thus, have been able to move on to the mature state of grammar as far as gender agreement is concerned.

The research questions we pursue in this study are as follows:

1. Do L1 English L2 Spanish speakers use semantic gender for gender agreement? In other words, do they perform gender agreement more accurately and more quickly on nouns that have semantic gender?

2. Do L1 English L2 Spanish speakers use morphophonological cues for gender agreement? In other words, do they perform gender agreement more accurately and more quickly on transparent nouns?

3. Does proficiency modulate the extent to which L1 English L2 Spanish speakers utilize agreement cues? In other words, do high-proficiency L2 learners perform better than low-proficiency L2 learners on *opaque* nouns, thus demonstrating that they are able to utilize agreement cues for gender assignment?

With respect to the first research question, if semantic gender serves as a cue for gender assignment in L2, L1 English L2 Spanish speakers should perform gender agreement more accurately on

nouns with semantic gender. Given the inconclusive findings of the existing literature on the effect of semantic gender on gender agreement, results of the present study will provide an additional and possibly elucidating piece of evidence on this matter. Moreover, the results will also suggest the role that semantic gender plays on gender assignment and, consequently, on gender agreement.

With respect to the second research question, finding that L1 English L2 Spanish speakers perform gender agreement more accurately and more quickly on transparent nouns would not only replicate the results of previous studies but would also provide additional evidence in favor of lexical explanations of L2 gender agreement difficulties. More specifically, less accurate performance on opaque gender nouns would indicate that the speaker is less able to effectively use the lexical route because she has not assigned all nouns of the L2 to a target-like gender class, and thus mostly relies on the form-based route, which represents a lexical, not a representational, issue.

With respect to the third research question, finding that high proficiency learners, unlike low proficiency learners, are capable of performing gender agreement accurately on opaque nouns would signal that L2 speakers' limited ability to employ agreement cues as reported in the eye-tracking literature is a temporary rather than a permanent phenomenon in the development of L2 grammars. Such a finding would provide additional evidence in favor of non-deficit accounts.

3.2 Method

3.2.1 Participants

Two groups of L1 English L2 Spanish learners (17 low proficiency and 16 high proficiency) and a control group of 17 native speakers of Spanish took part in the experiment. Their age and proficiency scores are summarized in **Table 1**. Participants in the two experimental groups were L1 English speakers who had started learning Spanish around puberty (10–14 years old). On the one hand, the low proficiency group was mainly composed of second and third semester Spanish students, who mostly had to take Spanish as a language requirement. On the other hand, the high proficiency group consisted of Spanish majors and graduate students at the Department

	n. of participants (total # = 50)	Age (years)	Proficiency score
English low proficiency	17	18–29	18–30
English high proficiency	16	20–35	40–49
Control group	17	26–42	n/a

Table 1: Age and proficiency scores.

of Spanish and Portuguese at Rutgers University. These high proficiency participants had lived in Spain and Spanish-speaking countries in Latin America and spoke Spanish daily in formal and informal communicative contexts. Thus, they were highly motivated, highly proficient and acculturated L2 speakers.

Proficiency levels in Spanish of the two experimental groups was measured via the written portion of a version of the Diploma de Español como Lengua Extranjera (DELE) (Embajada de España, Washington, DC), which is a typical measure of proficiency in Spanish administered by the Instituto Cervantes on behalf of the Spanish Ministry of Education and Science. The written portion of the test is designed to measure grammar and reading and consists of 50 total possible points (30 multiple choice answers and 20 cloze questions) (see Appendix). This study differentiated proficiency groups according to their score: the low-proficiency group scored between 18 and 30 while the high-proficiency group scored between 40 and 49. An independent samples t-test revealed significant differences between the low-proficiency group ($M = 26.06$, $SD = 3.944$) and high-proficiency group ($M = 43.50$, $SD = 3.141$), $t(31) = -13.993$, $p = .001$. Considering that all other variables were equal, these results confirm that the two experimental groups had different proficiency levels and could be used to test the research hypotheses.

The control group was composed of 17 native speakers of Spanish from Spain and from various countries in South and Central America, ages 26–42. All of them were undergraduate or graduate students at Rutgers, most of them in the Department of Spanish and Portuguese. Given the different origins of the control group participants, the items used in the tasks were not subject to dialectal variation in gender agreement patterns.

3.2.2 Tasks

The research task employed in the study was a self-paced reading grammaticality judgment task (GJT). It was created and executed on a laptop computer using PsychoPy software, and was presented in a center non-cumulative moving window format. While the validity of *the classical version of GJT* is often questioned for tapping into explicit, metalinguistic knowledge about the structure in question and not into implicit knowledge of the structure*, timed GJT* is considered by many researchers to be an appropriate tool to measure implicit knowledge (Bialystok 1979; Ellis et al. 2009; Bowles 2011; Rebuschat 2013; Godfroid et al. 2015; Spada et al. 2015; Zhang 2015).

In addition to using the timed version of the GJT, we created other conditions that encouraged the participants to use their implicit knowledge:

- the participants were asked to give their answers intuitively, as soon as possible and without thinking and were told that their reaction times would be recorded.
- the sentences were followed by a comprehension check, to encourage the participants to focus on meaning and not on grammar.

- since the sentences appeared on the screen one word at a time, the participants had to direct all of their attention to reading the words and had to hold the part of the sentence they had read in working memory. Therefore, it is highly likely that this task format imposed an increased processing load and prevented L2 learners from being able to retrieve their memorized knowledge of L2 rules.

- because of the non-cumulative presentation of the sentences the participants were not able to regress (move their gaze from right to left to see the previous words in the sentence), which is one of the ways in which L2 learners can reflect on their metalinguistic knowledge when they are taking a GJT (Godfroid et al. 2015).

- participants were naïve to the purposes of the experiment, and even after completion of it remained unaware of the linguistic structure in question (gender agreement).

Thus, with the understanding that this was a behavioral study, every effort was made to prevent participants from using metalinguistic knowledge of gender.

3.2.3 Stimuli

Test items consisted of 60 target sentences and 60 distractor sentences. The target sentences contained DPs with determiner/noun sequences, half of which (30) had gender agreement violations (were ungrammatical). The distractor sentences involved other types of violations such as misplaced prepositions and tense and number agreement violations. The sentences were 10–12 words long and constructed in such a way that the critical region (the DP) consistently appeared in the middle position.

The experimental DPs involved the following three conditions: a) a semantic gender condition with 20 animate nouns such as *abuela* 'grandparent-F', as in (17), b) an arbitrary gender condition with 20 inanimate nouns with transparent morphophonological cues (-*a* vs. -*o*) such as *cerveza* 'beer-F', as in (18) and c) an arbitrary gender condition with 20 nouns with opaque morpho-phonological cues such as *ley* 'law.F', as in (19). Nouns with semantic gender were morphologically transparent to match the nouns without semantic gender, and to ensure that the effect (if any) is caused by semantic gender per se. All experimental nouns were controlled for frequency – there were no statistically significant differences between the frequency of the nouns in grammatical vs. ungrammatical condition ($p = 0.551$), in animate vs. inanimate condition ($p = .072$), and in transparent vs. opaque condition ($p = .429$). Only nouns that are typically taught during the first two semesters of Spanish, and thus should be familiar to the participants, were used in the study.

(17) Semantic gender nouns (animate reference):

   *Me preocupa que <u>el</u>   <u>abuel-a</u>   se haya quedado sola en casa.
   CL worries that DEF.M grandparent-F CL has stayed alone in home
   'I am worried that the grandmother stayed at home alone.'

(18) Arbitrary gender nouns (inanimate reference) with transparent gender:

    *No me gusta <u>el</u>   <u>cervez-a</u> porque tiene un sabor amargo.

    not CL like DEF.M beer-F   because has a   taste bitter

    'I don't like beer because it has a bitter taste.'

(19) Arbitrary gender nouns (inanimate reference) with opaque gender:

    *En Estados Unidos el   ley   puede cambiar según   el   estado.

    in United States DEF.M law.F can   change according to the state

    'In United States the law can change according to the state.'

Due to technological issues during the execution of the experiment, seven of the target sentences were excluded from the analysis; hence, only fifty-three target sentences were analyzed. Out of those fifty-three sentences, seventeen contained semantic gender nouns, eighteen contained transparent gender nouns (seventeen of which also served as arbitrary gender nouns), and the remaining eighteen sentences contained opaque gender nouns.

3.2.4 Procedure

After reading detailed instructions on the computer screen and taking a practice test, the participants saw sentences in a center non-cumulative moving window format: they saw one word of the sentence at a time and every time they pressed the relevant key, the word disappeared and was followed by the next word of the sentence. After every word the participants saw, they had to indicate whether it looked acceptable in the context of the sentence.

The critical region was after the noun, because it is the region where participants had to judge whether the noun that they saw at the moment "agreed" (had a gender that matched the gender of the determiner) with the determiner that had preceded it. For example, the experimental sentence in (17) contains a DP with a gender agreement violation (*el abuel-a* 'DEF.M grandmother-F'). In this case the participants were expected to judge *abuela* (F) as unacceptable, because they had seen the determiner *el* 'DEF.M', whose gender value did not match that of the noun *abuela*. They got one (1) point for accuracy every time they correctly accepted a grammatical DP like *la abuela*, and every time they rejected an ungrammatical DP like *el abuela.

Our purpose was to tap into the participant's implicit knowledge of grammatical gender rather than to check their explicit knowledge of the rule. One of the ways to fulfill this purpose is to have participants focus on the meaning of experimental stimuli rather than on grammar. For that reason, all of the task sentences were followed by a comprehension check – a translation of the original Spanish sentence into English. The participants had to indicate by pressing a relevant key whether the translation corresponded to the original. The comprehension check translations matched the original sentences in Spanish only half of the time both for experimental and distractor sentences. Data from two low proficiency participants were discarded because

they answered comprehension questions correctly less than 80% of the time, which reduced the number of low proficiency participants from the initial 19 to 17.

We tested the effect (if any) of transparency by comparing DPs with nouns having transparent gender morphemes such as in (18) with DPs with nouns having opaque gender morphemes such as in (19), and the effect of semantic gender by comparing DPs with nouns having semantic gender such as in (17) with DPs with nouns having no semantic gender such as in (18). In addition, we tested whether the effects of the linguistic variables were modulated by proficiency. We were particularly interested in differences (if any) between the performance of the low proficiency and high proficiency groups on the opaque nouns, because such differences would be informative with respect to the groups' ability to use agreement cues and, therefore, would allow us to make inferences about the lexical vs. representational origin of L2 gender difficulties.

The RTs were recorded and analyzed to reveal any potential differences in the processing of grammatical vs. ungrammatical items, semantic vs. grammatical gender items, and transparent vs. opaque items.

To summarize, the study design in the experiment was 2 (Gender Type: semantic vs. arbitrary) × 2 (Transparency: transparent vs. opaque) × 2 (Grammaticality: grammatical vs. ungrammatical) × 3 (Group: low proficiency L2 vs. high proficiency L2 vs. controls). Accuracy and Reaction Times were the dependent variables.

4 Results
4.1 Accuracy

Table 2 presents the distribution of scores across the two experimental groups (low and high proficiency) and the control group for, on the one hand, the semantic gender condition, the

	Low proficiency	High proficiency	Controls
	M/total possible score	*M/total possible score*	*M/total possible score*
Semantic	14.44/17 (85%)	16.38/17 (96%)	16.65/17 (98%)
Arbitrary	14.39/17 (85%)	16.37/17 (96%)	16.94/17 (100%)
Transparent	15.39/18 (86%)	17.37/18 (97%)	17.94/18 (100%)
Opaque	10.94/18 (61%)	16.44/18 (91%)	17.29/18 (96%)
Grammatical	22.47/24 (94%)	23.81/24 (99%)	23.82/24 (99%)
Ungrammatical	13.82/24 (58%)	21.5/24 (90%)	23.41/24 (98%)

Table 2: Group mean scores and total possible scores (indicated next to the means).

arbitrary gender condition, the transparent morphology condition and the opaque morphology condition, and, on the other hand, for grammatical and ungrammatical items.

Accuracy was high for all three groups in most categories with two notable exceptions: low proficiency L2 speakers showed lower accuracy on opaque gender items and ungrammatical items.

4.1.1 Statistical analysis

In all models reported below, statistical significance of the variables and their interactions was assessed using hierarchical partitioning of variance via nested model comparisons.

The means in **Table 2** imply that the lower proficiency group had slightly lower accuracy across the board, but with substantially lower accuracy for ungrammatical items. To test these impressions, we fit a series of regression models (Lme4 package 1.1-21 in R 3.4.3). The results for the first model, with *Accuracy* as a function of *Group* and *Gender* (semantic vs. arbitrary) are presented in **Table 3**. Accuracy was significantly lower for the low proficiency group compared to the high proficiency group. The control group had higher accuracy than the high proficiency group, although this difference was not significant. Accuracy for semantic gender nouns was slightly but not significantly lower. In other words, the high proficiency and control groups seemed to pattern similarly when compared to the low proficiency participants in the conditions that contrasted semantic vs. arbitrary gender. However, the presence of a semantic gender on the noun had no significant effect on accuracy in any of the groups.

The results of the second model with *Accuracy* as response variable and *Group* and *Transparency* (transparent vs. opaque) as predictors are presented in **Table 4**.[3] Once again, low proficiency participants had significantly lower levels of accuracy compared to high proficiency ones, while the difference between the control group and the high proficiency group was not significant. Transparency did not have a significant effect on the high proficiency and control groups, but it

	Estimate	Std. Error	Z	*p*
Intercept	96.30	2.03	47.36	<2e–16
Low prof.	–11.08	2.43	–4.54	<1.57e–05
Control	2.51	2.47	1.01	0.31
Semantic	–0.11	1.98	–0.05	0.95
Adjusted R^2: 0.24				

Table 3: Results for Accuracy as a function of Group and Gender Type (semantic vs. arbitrary).

---

[3] This model only included nouns with inanimate references.

did have an effect on the low proficiency group. In other words, the high proficiency and control groups did not behave differently on transparent items compared to the opaque ones, but the low proficiency group was significantly more accurate on transparent items than on opaque ones. In sum, high proficiency speakers seem to pattern similarly to the controls, and transparency only had a clear effect on low proficiency participants.

The third model tested the effect of *Group* and *Grammaticality* on *Accuracy*, as seen in **Table 5**. The low proficiency group had lower accuracy than the high proficiency one, but not significantly so; similarly, the control group had slightly higher accuracy rates than the high proficiency participants, although this difference was also not significant. Ungrammatical items were rated significantly less

	Estimate	Std. Error	Z	*p*
Intercept	91.18	2.54	35.85	< 2e–16
Low prof.	–29.13	3.49	–8.33	< 5.53e–13
Control	4.75	3.54	1.34	0.18
Transparent	5.25	3.59	1.46	0.14
Low prof. × Transparent	18.08	4.94	3.65	0.0004
Control. × Transparent	–1.54	5.01	–0.30	0.75
Adjusted R^2: 0.61				

Table 4: Results for Accuracy as a function of morphological transparency (inanimate nouns only).

	Estimate	Std. Error	Z	*p*
Intercept	99.25	2.77	35.86	< 2e–16
Low prof.	–5.52	3.80	–1.45	0.15
Control	0.04	3.85	0.01	0.99
Ungrammatical	–9.50	3.91	–2.42	0.01
Low prof. × Ungrammatical	–26.50	5.38	–4.92	< 3.56e–09
Control × Ungrammatical	7.85	5.45	1.43	0.15
Adjusted R^2: 0.64				

Table 5: Results for Accuracy as a function of Grammaticality and Group.

accurately than grammatical ones, and this was particularly true for the low proficiency group, which showed a significant interaction with grammaticality. Together, these results suggest that the three groups make a clear distinction between grammatical and ungrammatical items.

4.2 Reaction times

In this section, we present results for reaction times. **Table 6** presents the distribution of mean RTs in milliseconds for the semantic gender condition, the arbitrary gender condition, the transparent and opaque morphology condition, and the grammatical vs. ungrammatical condition across the two experimental groups (low and high proficiency) and the control group.

4.2.2 Statistical analyses

For all of the experiments in this section, the response variable was *Reaction Time*. **Table 7** presents results for the model with RT as a response variable and *Group* and *Gender Type* (semantic vs. arbitrary) as predictors. The low proficiency group had significantly slower reaction times than the high proficiency group, whereas the control group had non-significantly faster reaction times than the high proficiency participants. Reaction times for nouns with semantic gender were slightly slower than for nouns with arbitrary gender, although not significantly so.

Coefficients for the model that assessed reaction times as a function of *Transparency* (transparent vs. opaque) and *Group* are presented in **Table 8**. As in previous models, lower proficiency participants had significantly longer reaction times compared to high proficiency participants; the control group had slightly shorter reaction times compared to high proficiency participants, but this difference was not significant. Finally, reaction times on transparent items were also significantly shorter than those on opaque ones.

	Mean RTs		
	Low proficiency	**High proficiency**	**Controls**
Semantic	1.380116	1.068312	1.009030
Arbitrary	1.324136	0.993908	0.990175
Transparent	1.303367	0.995254	0.977809
Opaque	1.494842	1.146679	0.986571
Grammatical	1.155087	0.905548	0.806824
Ungrammatical	1.694924	1.260203	1.201058

Table 6: Group mean RTs in milliseconds.

The results of the statistical model for reaction times as a function of *Grammaticality* (grammatical vs. ungrammatical) and *Group* are presented in **Table 9**. In this case, the low proficiency participants had significantly longer reaction times compared to the high proficiency

	Estimate	Std. Error	Z	*p*
Intercept	1.00	0.05	17.78	<2e–16
Low prof.	0.32	0.06	4.73	<7.53e–06
Control	–0.03	0.06	–0.45	0.64
Semantic	0.04	0.05	0.89	0.37
Adjusted R²: 0.24				

Table 7: Results for reaction times as a function of Gender Type (semantic vs. arbitrary) and Group.

	Estimate	Std. Error	Z	*p*
Intercept	1.13	0.06	18.23	<2e–16
Low prof.	0.32	0.07	4.41	<2.58e–05
Control	–0.08	0.07	–1.17	0.24
Transparent	–0.11	0.06	–1.94	0.05
Adjusted R²: 0.2682				

Table 8: Results for reaction times as a function of morphological Transparency and Group.

	Estimate	Std. Error	Z	*p*
Intercept	0.86	0.06	13.26	<2e–16
Low prof.	0.34	0.07	4.37	<3.09e–05
Control	–0.07	0.07	–0.99	0.32
Ungrammatical	0.43	0.06	6.78	<8.73e–10
Adjusted R²: 0.4324				

Table 9: Results for reaction times as a function of Grammaticality and Group.

group, whereas the control group had slightly shorter reaction times compared to the high proficiency participants, although this difference was not significant. Additionally, much longer reaction times were observed on ungrammatical items compared to grammatical ones, but there was no statistically significant interaction between *Group* and *Grammaticality*. Importantly, these results show that all three groups slowed down when they were presented with a gender agreement violation, which indicates that they did process those violations.

5 Discussion

Previous research has provided data to support non-deficit approaches by showing that variability in L2 morphosyntax can be caused by factors other than the absence of syntactic features, for example, by a complex process of feature reassembly (Lardiere 2009), by difficulties with realization of surface morphology (MSIF), by underusing existing syntactic knowledge (SSH), and by non-linguistic factors such as working memory capacity and decoding ability (McDonald 2006). Recently, studies on grammatical gender have found that some of the persistent L2 gender agreement errors are likely to be the result of underlying gender assignment errors. In other words, a *syntactic* issue (gender agreement) may in fact be caused by *lexical* factors (assignment). The present study elaborates on this finding by systematically exploring different gender assignment cues (animacy of the noun, morphophonological form, and gender of the agreeing elements), and by pinpointing how using or failing to use these assignment cues affects gender agreement in L2 learners.

This study was designed to address three research questions raised in section 3.1. With respect to the first research question, we found that semantic gender did not affect accuracy scores or reaction times in either of the L2 groups or in the control group. Some previous studies showed that semantic gender did facilitate L2 gender agreement (Finnemann 1992; Fernández-García 1999; Alarcón 2009), while other studies showed an opposite effect (Bruhn de Garavito & White 2002; Sagarra & Herschenson 2011; Sagarra & Herschenson 2012) and concluded that activating nouns with female semantic gender will cause activation of the male counterpart (and vice-versa) (e.g. *chico* 'boy activating *chica* 'girl'), thereby leading to lexical competition and a higher processing cost. The present finding, whereby there was no effect on accuracy, may be due to a canceling out of the facilitating effect of semantic gender and the confounding effect of increased processing cost of nouns with semantic gender.

With respect to the second research question, as hypothesized and in accordance with the existing literature, we found a robust effect of transparency on gender agreement in the low proficiency group: low proficiency L1 English L2 Spanish speakers performed gender agreement more accurately and more quickly on transparent nouns than on opaque nouns, revealing their ability to exploit morphophonological cues for gender agreement. For example, high proficiency

learners performed correctly 97% of the time – an almost perfect score – on transparent nouns (17.4 out of 18), and even low proficiency learners performed correctly 86% of the time (15.4 out of 18). Performance on opaque nouns in low proficiency group, on the other hand, is barely above chance (10.94/18 = 61%), which shows that low-proficiency speakers have no clear idea about the gender of many opaque nouns – a finding consistent with the notion that low proficiency speakers have not built abstract lexical representations of gender for L2 opaque nouns and thus cannot use the lexical route to access the gender of such nouns.

On the other hand, and interestingly enough, some of the low proficiency learners did not merely accept grammatical DPs such as *la casa* 'DEF.F house-F', but they also systematically rejected DPs with opaque nouns such as *la luz* 'DEF.F light.F' and *la miel* 'DEF.F honey.F', although these are grammatical in the target language (Spanish). This seems to indicate that they mark consonant endings such as *–z* and *–l* as masculine in their grammars – **el luz* DEF.M light.M' and **el miel* 'DEF.M honey.M'. This is a clear case of a gender assignment mismatch, similar to those made by L2 learners in Grüter et al. (2012). That is, low proficiency learners may default to assigning nouns with infrequent endings such as *luz* 'light' and *miel* 'honey' to the masculine class, making it a default class (Harris 1991). In any case, similarly to what Bruhn de Garavito & White (2002) suggest, these low proficiency learners seem to have assigned some opaque L2 nouns to distinct gender classes, and while these classes may be unstable and not target-like, they abide by the laws of the learners' developing grammars.

With respect to the third research question, high proficiency learners patterned significantly differently from low proficiency learners, and very close to the control group, confirming a clear acquisition progression and the idea that high proficiency learners do converge with native speakers on most aspects of gender assignment and agreement. Again, if we compare results from the two L2 groups in the different conditions, we note a clear trend: accuracy for morphologically transparent nouns increases from 86% for low proficiency group to 97% for high proficiency group, but even more importantly, for morphologically opaque nouns it changes from 61% (low proficiency) to 91% (high proficiency). Clearly, the wide gap between the scores on transparent and opaque nouns (86% vs. 61%) in the low proficiency group is considerably reduced in the high proficiency group (97% vs. 91%), showing that high proficiency learners are capable of using agreement cues to assign gender to opaque nouns. These results indicate that low-proficiency L2 learners mostly perform gender agreement superficially in the vocabulary component based on the morphophonological information provided by the noun, but that high-proficiency L2 learners do shift from the morphophonology-mediated agreement to the syntactic agreement based on the gender features of the agreeing elements (e.g. determiners). Thus, we find that the observed L2 overreliance on form-based (morphophonological) cues represents a developmental stage rather than a permanent state of L2 grammars.

The qualitative differences between our low and high proficiency L2 speakers with respect to gender agreement performance can be explained by a non-deficit representational account such as FRH (Lardiere 2008; 2009). Specifically, our L1 English L2 Spanish speakers, who have abstract gender features in their L1 system as part of personal pronouns (he/she) and some individual nouns such as 'ship' (she), need to map these features to a different system that marks gender on all nouns through transparent or opaque morphology. Such remapping is challenging, and convergence on the target will be harder when more remapping is required. Spinner (2013), for example, found that English learners had difficulty fully acquiring number in L2 Swahili because number is bundled with gender in the L2 as a prefix. Thus, L2 speakers must reassemble the feature *number* in two ways: as part of a prefix and as bundled with the new gender feature. Guijarro-Fuentes (2014) also analyzes the Spanish L2 non-target gender and number agreement patterns of French L1 learners in terms of a difficulty in reassembling the bundles from the L1 to the L2. Cuza & Pérez-Tattam (2016) found varying degrees of agreement in the DP of the Spanish of bilingual heritage speakers in contact with English. Their account proposes that the [+strong] feature in the Spanish DP needs to be remapped to English's [–strong] feature. Similarly, our low proficiency participants' performance indicates an initial failure to map L1 gender forms (or lack thereof) to the L2 gender forms. At the same time, one of the fundamental predictions of the FRH is that target-like L2 acquisition is possible, and that after initial failure to map these features one-to-one, a slow and gradual reorganization process will take place in the L2 learners' developing grammatical system, eventually leading to (potentially) full convergence between the L1 and L2 systems. This prediction is consistent with our findings in high proficiency group: despite their L1 (English) lacking grammatical gender and agreement, they were able to shift to syntactic agreement after sufficient experience with the target system.

In addition, since gender assignment is to a great extent a lexical process, all factors that contribute to lexical learning are also relevant for gender assignment. Variable amount of exposure to the L2 is probably the most influential lexical factor to account for the differences between our low and high proficiency participants' performance on opaque nouns (Unsworth 2008; Putnam & Sánchez 2013). Specifically, since gender assignment involves learning language-specific gender values for each individual noun of a given language, it represents a lengthy, incremental process that depends on the amount of exposure and practice and that may take years and even decades to complete. Low proficiency learners have had shorter lengths of exposure to the L2 than high proficiency learners and thus had less time to memorize the gender of opaque gender nouns; hence, our finding that low proficiency learners were less accurate on such nouns was only to be expected. Moreover, input alone is not sufficient unless it becomes intake, which can happen only via processing for comprehension and for production (Putnam & Sánchez 2013). Obviously, our low proficiency L2 speakers, who are second and third semester L2 learners, have had fewer opportunities to comprehend and produce the L2 compared to our high proficiency L2 speakers,

who are Spanish majors or graduate students who have lived in Spanish-speaking countries and interacted with native speakers on a daily basis.

The difficulties with opaque gender nouns observed in low proficiency learners are also compatible with the learning/teaching context hypothesis (Arnon & Ramscar 2012; Grüter et al. 2012; Hopp 2015). First, adult L2 speakers typically learn their L2 in a classroom, where teaching grammatical gender focuses on and draws learners' attention to noun endings. Second, exposure to the L2 in a classroom to a great extent happens via written input, where phrases and sentences contain L2 words clearly separated by spaces (Grüter et al. 2012). This is very different from exposure to L1, which happens exclusively via unsegmented aural input that contains strings of sounds where the determiner and the noun occur together reinforcing the link between the gender of the noun and the determiner. Exposure to written input allows L2 learners to see words separately and immediately assign semantic meanings to content words (e.g. nouns, verbs, etc.), but it also increases their lack of attention to morphosyntactic information marked on (gender) morphemes. Conversely, exposure to aural input forces child L1 learners to focus on transitional probabilities to figure out where word boundaries are, and only then assign meanings to words.

In fact, beginning L2 learners have been shown to display a general tendency to process semantic and pragmatic information before attending to the morphosyntactic cues (Giacalone Ramat & Banfi 1990; Dietrich et al. 1995; Bardovi-Harlig 1999; Clahsen & Felser 2006). For example, when processing tense, they tend to attend to semantic information through lexical items such as adverbs, not morphosyntactic features like morphemes or agreement. It may be the case that the two factors – exposure to written input instead of unsegmented aural input (learning experience) and the L2 tendency to prioritize semantic and pragmatic cues (learning type) – conspire against *beginning* L2 learners, hindering them from attending to agreement cues, which are crucial for assigning gender to opaque nouns. Hopp (2015: 477) expressed a similar idea, "Hence, early learning of the direct mapping from nouns to meaning blocks learning of the agreement association between article and noun, so that differences in the learning type and the learning experience may account for why children and adults differ in their reliance on predictive processing, even if the amount of exposure to the relevant distributional cooccurrences is (roughly) equal."

Low proficiency participants in this study were second or third semester students in Spanish classes who had learned gender in the typical classroom way – by being explicitly taught that nouns that end in –*a* are feminine and those that end in –*o* are masculine, and their exposure to the L2 was at least half the time via written input. High proficiency participants, on the other hand, were mostly Spanish majors and graduate students in the Department of Spanish and Portuguese, who had been exposed to numerous hours of oral input through, among other things, interacting with their native-speaking peers in the Department or having lived abroad.

Thus, the findings of the present study add more data to the growing body of research on how variability in lexical gender assignment affects grammatical gender acquisition in L2. Most

importantly, we approach gender assignment by analyzing the different assignment cues and show how using or failing to use those cues affects L2 gender agreement.

6 Limitations

As with any study, this work has limitations. First, we have used only one experiment, although we did examine both accuracy and RTs. Second, the number of participants per group is lower than in some other studies. This latter limitation may have led to the lack of significant differences on RTs between our high-proficiency group and the control group, despite, as an anonymous reviewer notes, the normal tendency for even highly proficient adult L2 learners to have slower RTs. Importantly, these limitations have not prevented us from being able to test our research hypotheses and to provide more data to the growing body of research on lexical approaches to L2 morphosyntactic variability as well as on non-deficit explanations of L2 gender agreement errors.

7 Conclusions

This paper makes a fundamental assumption that gender agreement is not the same as gender assignment. Gender assignment has to do with assigning nouns of a gendered language to two or more gender classes in the learner's developing lexicon – grammatical gender, while gender agreement has to do with sharing the gender of a given noun with all of the agreeing elements in the syntactic structure. Language learners are capable of assigning gender to nouns because they have the ability to track gender morphemes and process gender agreement in the input they are exposed to.

The main finding of this study is that the difficulties usually associated with gender agreement are in fact often caused by gender assignment issues. These difficulties may be easily confused for one another because the two – gender assignment and gender agreement – are tightly interconnected: gender agreement cannot be performed if the noun has not been assigned to a gender class, and gender cannot be assigned to the target-like gender class unless agreement cues are attended to in the input. Low proficiency L2 learners rely on morphophonological cues, but these cues are not entirely reliable and, most importantly, are not available on opaque nouns. This leads them to non-target-like gender assignment, which in turn causes non-target like gender agreement. At the same time, when morphophonological cues *are* available, even low proficiency learners perform fairly accurately on gender agreement.

Thus, gender agreement deficits can be at least partially attributed to (over)reliance on morphophonological cues accompanied by a limited ability to exploit agreement cues in the input while assigning gender to nouns. Crucially, based on the results of the present study, this deficit is only temporary, and is overcome with higher proficiency and more exposure to L2 input.

Abbreviations

DEF = definite, CL = clitic, M = masculine, F = feminine, N = noun

Additional File

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

- **Supplementary file 1: Appendix.** DELE test. DOI: https://doi.org/10.16995/glossa.5761.s1

Acknowledgements

Competing Interests

References

Akhutina, Tatiana & Kurgansky, Andrei & Polinsky, Maria & Bates, Elizabeth. 1999. Processing of grammatical gender in a three-gender system: Experimental evidence from Russian. *Journal of Psycholinguistic Research* 28(6). 695–713.

Alarcón, Irma. 2004. The sequential acquisition of L2 Spanish gender marking: Assignment and agreement. In José, Brian & de Jong, Ken (eds.), *IULC Working Papers* 4(1), article 7. (https://scholarworks.iu.edu/journals/index.php/iulcwp/article/view/25960). (Accessed 2021-09-25).

Alarcón, Irma. 2006. *The second language acquisition of Spanish gender agreement: The effects of linguistic variables on accuracy*. Munich: Lincom Europa.

Alarcón, Irma. 2009. The processing of gender agreement in L1 and L2 Spanish: Evidence from reaction time data. *Hispania* 94(4). 814–828.

Alarcón, Irma. 2011. Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition* 14. 332–350. DOI: https://doi.org/10.1017/S1366728910000222

Arnon, Inbal & Ramscar, Michael. 2012. Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition* 122(3). 292–305. DOI: https://doi.org/10.1016/j.cognition.2011.10.009

Bardovi-Harlig, Kathleen. 1999. From morpheme studies to temporal semantics: Tense-aspect research in SLA. *Studies in Second Language Acquisition* 21(3). 341–382. DOI: https://doi.org/10.1017/S0272263199003010

Bates, Elizabeth & Devescovi, Antonella & Pizzamiglio, Luigi & D'Amico, Simona & Hernandez, Arturo. 1995. Gender and lexical access in Italian. *Perception and Psychophysics* 57. 847–862. DOI: https://doi.org/10.3758/BF03206800

Bialystok, Ellen. 1979. Explicit and implicit judgments of L2 grammaticality. *Language Learning* 29. 81–103. DOI: https://doi.org/10.1111/j.1467-1770.1979.tb01053.x

Bordag, Denisa. 2004. Interaction of L1 and L2 systems at the level of grammatical encoding: Evidence from picture naming. In Foster-Cohen, Susan (ed.), *EUROSLA Yearbook 4*, 203–230. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/eurosla.4.10int

Bordag, Denisa & Opitz, Andreas & Pechmann, Thomas. 2006. Gender processing in first and second languages: The role of noun termination. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32(5). 1090–1101. DOI: https://doi.org/10.1037/0278-7393.32.5.1090

Bowles, Melissa. 2011. Measuring implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition* 33. 247–271. DOI: https://doi.org/10.1017/S0272263110000756

Bruhn de Garavito, Joyce & White, Lydia. 2002. L2 acquisition of Spanish DPs: The status of grammatical features. In Pérez-Leroux, Ana Teresa & Liceras, Juana M. (eds.), *The acquisition of Spanish morphosyntax: The L1/L2 connection*, 153–178. Dordrecht, the Neverlands: Kluwer. DOI: https://doi.org/10.1007/978-94-010-0291-2_6

Caffarra, Sendy & Janssen, Niels & Barber, Horacio H. 2014. Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia* 63. 124–134. DOI: https://doi.org/10.1016/j.neuropsychologia.2014.08.016

Cain, Jacqueline & Weber-Olsen, Marcia & Smith, Rosslyn. 1987. Acquisition strategies in a first and second language: Are they the same? *Journal of Child Language* 14(2). 333–352. DOI: https://doi.org/10.1017/S0305000900012964

Caramazza, Alfonso. 1997. How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 14. 177–208. DOI: https://doi.org/10.1080/026432997381664

Carroll, Susanne. 1989. Second language acquisition and the computational paradigm. *Language Learning* 39. 535–594. DOI: https://doi.org/10.1111/j.1467-1770.1989.tb00902.x

Carstens, Vicky. 2000. Concord in minimalist theory. *Linguistics Inquiry* 31. 319–355. DOI: https://doi.org/10.1162/002438900554370

Carstens, Vicky. 2010. Implications of grammatical gender for the theory of uninterpretable features. In Putnam, Michael (ed.), *Exploring Crash Proof Grammars*, 31–57. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/lfab.3.03car

Clahsen, Harald & Felser, Claudia. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27(1). 3–42. DOI: https://doi.org/10.1017/S0142716406060024

Cuza, Alejandro & Pérez-Tattam, Rocío. 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. DOI: https://doi.org/10.1017/S1366728914000893

Dahan, Delphine & Swingley, Daniel & Tanenhaus, Michael K. & Magnuson, James S. 2000. Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language* 42(4). 465–480. DOI: https://doi.org/10.1006/jmla.1999.2688

Davies, Mark. 2002. Corpus del Español: 100 Million Words, 1200s–1900s. Retrieved from http://www.corpusdelespanol.org/.

Dewaele, Jean-Marc & Véronique, Daniel. 2001. Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition* 4(3). 275–297. DOI: https://doi.org/10.1017/S136672890100044X

Dietrich, Rainer & Klein, Wolfgang & Noyau, Colette. 1995. *The acquisition of temporality in a second language*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/sibil.7

Dussias, Paola & Valdés Kroff, Jorge & Guzzardo Tamargo, Rosa & Gerfen, Chip. 2013. When gender and looking go hand in hand. *Studies in Second Language Acquisition* 35(2). 353–387. DOI: https://doi.org/10.1017/S0272263112000915

Ellis, Rod & Loewen, Shawn & Elder, Catherine & Reinders, Hayo & Erlam, Rosemary & Philp, Jenefer. 2009. *Implicit and explicit knowledge in second language learning, testing and teaching*. Tonawanda, NY: Multilingual Matters.

Fernández-García, Marisol. 1999. Patterns of gender agreement in the speech of second language learners. In Gutiérrez-Rexach, Javier & Martínez-Gil, Fernando (eds.), *Advances in Hispanic linguistics: Papers from the 2nd Hispanic linguistics symposium*, 3–15. Somerville, MA: Cascadilla Press.

Finnemann, Michael D. 1992. Learning agreement in the noun phrase: The strategies of three first-year Spanish students. *IRAL-International Review of Applied Linguistics in Language Teaching* 30(2). 121–136. DOI: https://doi.org/10.1515/iral.1992.30.2.121

Franceschina, Florencia. 2005. *Fossilized second language grammars: The acquisition of grammatical gender*. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/lald.38

Giacalone Ramat, Anna & Banfi, Emanuele. 1990. The acquisition of temporality. A second language perspective. *Folia Linguistica* 24(3–4). 405–428. DOI: https://doi.org/10.1075/sibil.7

Godfroid, Aline & Loewen, Shawn & Jung, Sehoon & Park, Ji-Hyun & Gass, Susan & Ellis, Rod. 2015. Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye movement patterns. *Studies in Second Language Acquisition* 37(2). 269–297. DOI: https://doi.org/10.1017/S0272263114000850

Gollan, Tamar & Frost, Ram. 2001. Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research* 30. 627–651. DOI: https://doi.org/10.1023/A:1014235223566

Grüter, Theres & Lew-Williams, Casey & Fernald, Anne. 2012. Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research* 28. 191–215. DOI: https://doi.org/10.1177/0267658312437990

Guijarro-Fuentes, Pedro. 2014. On the L2 Spanish acquisition of [+/−] interpretable features. In Costa, João & Fiéis, Alexandra & Freitas, Maria João & Lobo, Maria & Santos, Ana Lúcia (eds.), *New directions in the acquisition of Romance languages: Selected proceedings of the Romance Turn V*, 101–133. Newcastle upon Tyne: Cambridge Scholars Publishing.

Hammer, Anke & Goebel, Rainer & Schwarzbach, Jens & Munte, Thomas F. & Jansma, Bernadette M. 2007. When sex meets syntactic gender on a neural basis during pronoun processing. *Brain Research* 1146. 185–198. DOI: https://doi.org/10.1016/j.brainres.2006.06.110

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

Hawkins, Roger. 2009. Statistical learning and innate knowledge in the development of second language proficiency: Evidence from the acquisition of gender concord. In Benati, Alessandro G. (ed.), *Issues in second language proficiency*, 63–78. London: Continuum International Publishing.

Hawkins, Roger & Casillas, Gabriela. 2008. Explaining frequency of verb morphology in early L2 speech. *Lingua* 118(4). 595–612. DOI: https://doi.org/10.1016/j.lingua.2007.01.009

Hawkins, Roger & Chan, Cecilia Yuet-hun. 1997. The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research* 13. 187–226. DOI: https://doi.org/10.1191/026765897671476153

Hawkins, Roger & Franceschina, Florencia. 2004. Explaining the acquisition and non-acquisition of determiner–noun gender concord in French and Spanish. In Prévost, Philippe & Paradis, Johanne (eds.), *The acquisition of French in different contexts. Focus on functional categories*, 175–207. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/lald.32.10haw

Haznedar, Belma & Schwartz, Bonnie D. 1997. Are there Optional Infinitives in child L2 acquisition? In Hughes, Elizabeth & Hughes Mary & Greenhill, Annabel (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development*, 257–268. Somerville, MA: Cascadilla Press.

Heim, Stefan 2008. Syntactic gender processing in the human brain: A review and a model. *Brain and Language* 106. 55–64. DOI: https://doi.org/10.1016/j.bandl.2007.12.006

Heim, Stefan & Eickhoff, Simon B. & Opitz, Bertram & Friederici, Angela D. 2006. BA 44 in Broca's area supports syntactic gender decisions in language production. *Neuroreport* 17(11). 1097–1101. DOI: https://doi.org/10.1097/01.wnr.0000224775.05023.42

Hopp, Holger. 2013. Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research* 29(1). 33–56. DOI: https://doi.org/10.1177/0267658312461803

Hopp, Holger. 2014. Working memory effects in the L2 processing of ambiguous relative clauses. *Language Acquisition* 21(3). 250–278. DOI: https://doi.org/10.1080/10489223.2014.892943

Hopp, Holger. 2015. Differential learning of prediction. *Linguistic Approaches to Bilingualism* 5(4). 476–481. DOI: https://doi.org/10.1075/lab.5.4.06hop

Hopp, Holger. 2016. Learning (not) to predict: Grammatical gender processing in adult L2 acquisition. *Second Language Research* 32(2). 277–307. DOI: https://doi.org/10.1177/0267658315624960

Indefrey, Peter & Levelt, Willen J. M. 2004. The spatial and temporal signatures of word production components. *Cognition* 92(1–2). 101–144. DOI: https://doi.org/10.1016/j.cognition.2002.06.001

Johnson, Elizabeth K. 2005. Grammatical gender and early word recognition in Dutch. In Brugos, Alejna & Clark-Cotton, Manuella & Ha, Seungwan (eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development (BUCLD 29)*, 320–330. Somerville, MA: Cascadilla Press.

Karmiloff-Smith, Annette. 1979. *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.

Kupisch, Tania & Akpınar, Deniz & Stöhr, Antje. 2013. Gender assignment and gender agreement in adult bilinguals and second language learners of French. *Linguistic Approaches to Bilingualism* 3(2). 150–179. DOI: https://doi.org/10.1075/lab.3.2.02kup

Lardiere, Donna. 2008. Feature assembly in second language. In Liceras, Juana M. & Zobl, Helmut & Goodluck, Helen (eds.), *The role of formal features in second language acquisition*, 106–140. Mahwah, NJ: Lawrence Erlbaum. DOI: https://doi.org/10.4324/9781315085340

Lardiere, Donna. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25(2). 173–227. DOI: https://doi.org/10.1177/0267658308100283

Levelt, Willen. J. M. & Roelofs, Ardi & Meyer, Antje S. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22. 1–75. DOI: https://doi.org/10.1017/S0140525X99001776

Lew-Williams, Casey & Fernald, Anne. 2007. Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science* 18(3). 193–198. DOI: https://doi.org/10.1111/j.1467-9280.2007.01871.x

Lew-Williams, Casey & Fernald, Anne. 2010. Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language* 63(4). 447–464. DOI: https://doi.org/10.1016/j.jml.2010.07.003

Martin, Andrea E. & Monahan, Phillip J. & Samuel, Arthur G. 2017. Prediction of agreement and phonetic overlap shape sublexical identification. *Language and Speech* 60(3). 356–376. DOI: https://doi.org/10.1177/0023830916650714

McDonald, Janet L. 2006. Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language* 55. 381–401. DOI: https://doi.org/10.1016/j.jml.2006.06.006

Miceli, Gabriele & Turriziani, Patrizia & Caltagirone, Carlo & Capasso, Rita & Tomaiuolo, Francesco & Caramazza, Alfonso. 2002. The neural correlates of grammatical gender: An fMRI investigation. *Journal of Cognitive Neuroscience* 14(4). 618–628. DOI: https://doi.org/10.1162/08989290260045855

Montrul, Silvina & Foote, Rebecca & Perpinán, Silvia. 2008. Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning* 58(3). 503–553. DOI: https://doi.org/10.1111/j.1467-9922.2008.00449.x

Müller, Natascha. 1994. Gender and number agreement within DP. In Meisel, Jürgen M. (ed.), *Two first languages: Early grammatical development in bilingual children*, 53–88. Dordrecht: Foris. DOI: https://doi.org/10.1075/lald.7.05mul

Oliphant, Katrina. 1998. Acquisition of grammatical gender in Italian as a foreign language. *The Canadian Modern Language Review* 54(2). 239–262. DOI: https://doi.org/10.3138/cmlr.54.2.239

Padovani, Roberto & Calandra-Buonaura, Giovanna & Cacciari, Cristina & Benuzzi, Francesca & Nichelli, Paolo. 2005. Grammatical gender in the brain: Evidence from an fMRI study on Italian. *Brain Research Bulletin* 65(4). 301–308. DOI: https://doi.org/10.1016/j.brainresbull.2004.11.025

Pérez-Pereira, Miguel. 1991. The acquisition of gender: What Spanish children tell us. *Journal of Child Language* 18. 571–590. DOI: https://doi.org/10.1017/S0305000900011259

Pesetsky, David & Torrego, Esther. 2004. The syntax of valuation and the interpretability of features. In Karimi, Simin & Samiian, Vida & Wilkins, Wendy (eds.), *Phrasal and Clausal*

Architecture: Syntactic Derivation and Interpretation, 262–294. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.101.14pes

Prévost, Phillipe & White, Lydia. 2000. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16. 103–133. DOI: https://doi.org/10.1191/026765800677556046

Putnam, Michael & Sánchez, Liliana. 2013. What's so incomplete about incomplete acquisition?: A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism* 3(4). 478–504. DOI: https://doi.org/10.1075/lab.3.4.04put

Quiñones, Ileana & Molinaro, Nicola & Mancini, Simona & Hernández-Cabrera, Juan Andrés & Barber, Horacio & Carreiras, Manuel. 2018. Tracing the interplay between syntactic and lexical features: fMRI evidence from agreement comprehension. *NeuroImage* 175. 259–271. DOI: https://doi.org/10.1016/j.neuroimage.2018.03.069

Rebuschat, Patrick. 2013. Measuring implicit and explicit knowledge in second language research. *Language Learning* 63. 595–626. DOI: https://doi.org/10.1111/lang.12010

Sabourin, Laura & Stowe, Laurie A. 2008. Second language processing: When are L1 and L2 processed similarly. *Second Language Research* 24. 397–430. DOI: https://doi.org/10.1177/0267658308090186

Sagarra, Nuria & Herschensohn, Julia. 2011. Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning* 61(1). 80–116. DOI: https://doi.org/10.1111/j.1467-9922.2010.00588.x

Sagarra, Nuria & Herschensohn, Julia. 2012. Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism* 17(5). 607–627. DOI: https://doi.org/10.1177/1367006912453810

Schlig, Carmen. 2003. Analysis of agreement errors made by third-year students. *Hispania* 86(2). 312–319. DOI: https://doi.org/10.2307/20062864

Sekerina, Irina A. & Brooks, Patricia J. & Kempe, Vera. 2006. Gender transparency facilitates noun selection in Russian. In Lavine, James E. (ed.), *The proceedings of the 14th annual workshop on formal approaches to Slavic linguistics. The Princeton Meeting, 2005*. University of Michigan: Michigan Slavic Publications.

Smolík, Filip. 2014. The use of gender information in lexical processing in Czech 23-month-olds: An eye-tracking study. In Orman, Will & Valleau, Matthew James (eds.), *BUCLD 38: Proceedings of the 38th annual Boston University Conference on Language Development*, 451–462. Somerville, MA: Cascadilla Press.

Spada, Nina & Shiu, Julie Li-Ju & Tomita, Yasuyo. 2015. Validating an elicited imitation task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning* 65. 723–751. DOI: https://doi.org/10.1111/lang.12129

Spinner, Patti. 2013. The second language acquisition of number and gender in Swahili: A Feature Reassembly approach. *Second Language Research* 29(4). 455–479. DOI: https://doi.org/10.1177/0267658313477650

Spinner, Patti & Juffs, Alan. 2008. L2 grammatical gender in a complex morphological system: The case of German. *IRAL-International Review of Applied Linguistics in Language Teaching* 46(4). 315–348. DOI: https://doi.org/10.1515/IRAL.2008.014

Spinner, Patti & Thomas, Jamie A. 2014. L2 learners› sensitivity to semantic and morphophonological information on Swahili nouns. *International Review of Applied Linguistics in Language Teaching* 52(3). 283–311. DOI: https://doi.org/10.1515/iral-2014-0013

Taraban, Roman & Kempe, Vera. 1999. Gender processing in native and non-native Russian speakers. *Applied Psycholinguistics* 20. 119–148. DOI: https://doi.org/10.1017/S0142716499001046

Taraban, Roman & Roark, Bret. 1996. Competition in learning language-based categories. *Applied Psycholinguistics* 17. 125–148. DOI: https://doi.org/10.1017/S014271640000761X

Teschner, Richard V. & Russell, William M. 1984. The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics* 1(1). 115–132.

Tsimpli, Ianthi M. & Dimitrakopoulou, Maria. 2007. The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research* 23(2). 215–242. DOI: https://doi.org/10.1177/0267658307076546

Unsworth, Sharon. 2008. Age and input in the acquisition of grammatical gender in Dutch. *Second Language Research* 24(3). 365–395. DOI: https://doi.org/10.1177/0267658308090185

Van Heugten, Marieke & Johnson, Elizabeth K. 2011. Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. *Journal of Child Language* 38(1). 87–100. DOI: https://doi.org/10.1017/S0305000909990146

White, Lydia & Valenzuela, Elena & Kozlowska-Macregor, Martyna & Leung, Yan-Kit Ingrid. 2004. Gender and number agreement in non-native Spanish. *Applied Psycholinguistics* 25. 105–133. DOI: https://doi.org/10.1017/S0142716404001067

Zhang, Runhan. 2015. Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition* 37. 457–486. DOI: https://doi.org/10.1017/S0272263114000370