# Sensitivity to Gender Agreement in Second Language Hindi: A Processing Investigation

**Claire Renaud and Lauren Covey**
**Arizona State University**

## 1. Introduction

Recent developments in theories of language (i.e., the Minimalist Program; Chomsky, 1995, among others) focus on the level of features. There are two types of features: Interpretable features are arguably those that convey semantic information, whereas uninterpretable features are those that express grammatical information. For example, nouns carry the interpretable gender feature, whereas adjectives carry the uninterpretable gender feature used to compute agreement via the Agree operation. A current debate in adult second language (L2) acquisition focuses on whether a feature not present in the learner's first language (L1) grammar can be accessed in L2 acquisition (e.g., Hawkins & Chan, 1997; Lardiere, 2008, 2009; McCarthy, 2007, 2008; Prévost & White, 2000; Tsimpli & Dimitrakopoulou, 2007). In the present article, we investigate this debate more specifically by testing the predictions of the Interpretability Hypothesis (e.g., Tsimpli & Dimitrakopoulou, 2007), which argues that features will not be accessible in L2 acquisition, against those of the Feature Reassembly Hypothesis (e.g., Lardiere, 2009), which argues that features will be acquirable, but that relieving processing costs is crucial and reflects the hierarchy of the feature-organization (e.g., Harley & Ritter, 2002). To do so, we present the findings of a pilot study on overt adjective agreement in L1 English-L2 Hindi learners that were obtained with a self-paced reading judgment task. Because the uninterpretable gender feature is absent in English, but present in Hindi, this language pairing allowed for testing whether L2 learners rely on language-specific mechanisms—such as the Agree operation (see, e.g., Chomsky, 1995)—involving an uninterpretable feature absent from their L1 grammar during reading.

## 2. On the two hypotheses

### *2.1. Inaccessability of features*

Hawkins and Chan's (1997) Failed Functional Feature Hypothesis, later refined by Tsimpli and Dimitrakopoulou (2007) as the Interpretability Hypothesis, states that functional, uninterpretable features absent in the L1 grammar are subject to a critical period. This means that, past a certain age (usually taken to be around puberty), L2 learners who attempt to learn a language that instantiates different uninterpretable features will be unable to make use of these features in L2 acquisition, as they are confined only to the L1 uninterpretable features. Therefore, although the L2 grammar is believed to operate on UG principles, learners will resort to context-sensitive rules based on distributional regularities in the input (Hawkins & Casillas, 2008). In other words, learners will rely on frequency of occurrence in the input to achieve more automatized processing routines, but those rules will remain domain-general in nature (i.e., the L2 grammar will not be able to rely on language-specific mechanisms). Such rules will presumably consist of associating identified features on a matching versus mismatching basis.

* We would like to thank the audience from SLRF as well as the anonymous reviewers for their valuable feedback. Part of this project was supported by the Sun Angel Excellence in the Humanities Research Scholarship from the College of Liberal Arts and Sciences at Arizona State University, awarded to the second author.

 Selected Proceedings of the 2012 Second Language Research Forum, ed. Ryan T. Miller et al., 51-64. Somerville, MA: Cascadilla Proceedings Project.

Several studies provided support for this proposal (e.g., Franceschina, 2005; Hawkins & Casillas, 2008; Hawkins & Chan, 1997; Hawkins & Liszka, 2003). For instance, Franceschina (2005) investigated knowledge of gender assignment and agreement in two groups of near-native Spanish L2 learners: The [+gender] group ($n$ = 53) consisted of learners whose L1 also exhibited gender agreement (i.e., Arabic, French, German, Greek, Italian, and Portuguese) and the [-gender] group ($n$ = 15) included English learners, a language that does not exhibit grammatical gender. A control group of 42 native Spanish speakers was also included. Participants completed a variety of production and comprehension tasks. The [+gender] group and the native speakers always produced gender agreement correctly, whereas the L1 English speakers exhibited some mistakes (with around 90% accuracy). Despite this difference between the [-gender] and the [+gender] groups, the lowest accuracy score of the [-gender] group was 86.99% on gender agreement marked on pronouns. Throughout the tasks, a similar disadvantage for the [-gender] group was in evidence, whereas the [+gender] group patterned like the native speakers. It is interesting to note that, on the novel word task, even though all participants were found to rely on gender cues, the native speakers and the [+gender] group "relied on masculine clues more strongly than on feminine clues, while the opposite was true of the -gen[der] L1 group" (Franceschina, 2005, p. 180). Furthermore, all participants performed at ceiling on the gender assignment task, which indicates that they had knowledge of the gender of the nouns used in the study. Franceschina argues that these findings provide support for the Interpretability Hypothesis (e.g., Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007) because the [-gender] group seems to have acquired the interpretable gender feature, as evidenced by their excellent performance on the gender assignment task, but not the uninterpretable gender feature, as shown by the morphological variability and lack of accuracy when this feature is involved. There was no such asymmetry in the [+gender] learner group, which points to an advantage of having the feature instantiated in the L1 grammar.

### *2.2. Accessibility and hierarchy of features*

Another group of researchers argues that all features are instantiated in the L2 grammar, but that communication pressure and processing demands may yield the insertion of default (i.e., less specified) forms (e.g., Hazdenar & Schwartz, 1997; McCarthy, 2008; Prévost & White, 2000). Couched within the Minimalist framework, Lardiere refined these proposals with the Feature Reassembly Hypothesis (e.g., 2009). She proposed that all features can presumably be detected; however, L2 learners face the challenging task of reassembling (i.e., reorganizing) the relevant features into appropriate L2 representations, in part because of the influence of the feature organization of the L1 grammar. Because the task of (re)assembly necessitates significant resources, default forms may be inserted as a means to relieve processing costs (e.g., McCarthy, 2008; Prévost & White, 2000). Assuming Distributed Morphology (e.g., Halle & Marantz, 1993), Harley and Ritter (2002) proposed that (interpretable) features are modeled in hierarchical geometries. They argued that all languages have the feature of individuation (number). This feature contains the feature class, which further consists of animacy and gender. Importantly, such a feature hierarchy encodes markedness relations in that the more marked the feature is, the further embedded it is. For instance, gender is more marked than number because it requires the specification of both individuation and class, whereas number only requires for individuation to be specified. The notion of markedness implies underspecification, in which all redundant information is excluded from a representation. Although the nature of the processes whereby (new) features are selected and reassembled is not discussed by Lardiere, the manner in which features are computed might be highly revealing of the fine-grained nature of L2 grammars. Additionally, because we assume that what cannot be parsed cannot be learned (e.g., Fodor, 1998), grammatical development can only occur when a parse is required by the input but unlicensed by the current grammar. We further assume that the universal parser provides the essential learning mechanism that governs L2 acquisition (e.g., Dekydtspotter, 2001). In other words, to trigger a UG-constrained grammatical change, the parser needs to be minimally Universal Grammar (UG)-compatible. Even though the grammar may be target deviant at earlier stages of acquisition, representations are nonetheless expected to be computed by the parser by using aspects of the representations themselves (see Fodor, 1983). Learners are thus expected to rely on the information

provided by the input and on UG-constrained operations (such as Agree), which may reflect the organization of features in a hierarchical model.

In a study similar to Franceschina's (2005), McCarthy (2008) examined L2 Spanish learners' acquisition of gender and number agreement with clitics and adjectives. Twenty-four English learners of Spanish (15 intermediate and 9 advanced learners) and a group of 10 Spanish native speakers were asked to complete three tasks: an elicited production task, a picture identification task, and a vocabulary test. In the production task, participants were generally more accurate with number than with gender agreement, with the intermediate learners exhibiting at least 90% accuracy, except on plural agreement on adjectives (with 68.2% accuracy). Similarly, the advanced learners scored near ceiling on all agreement types except on plural agreement on adjectives (83.9% accuracy). Additionally, an ANOVA on gender agreement cases indicated greater accuracy overall for masculine than for feminine contexts ($p < .01$). The comprehension tasks yielded similar results: feminine clitics (intermediate learners 90.9%, advanced learners 98.4%) yielded higher accuracy scores than masculine clitics (intermediate learners 68.4%, advanced learners 96.3%). Crucially, although the intermediate learners only produced morphological agreement around 60% in some conditions, McCarthy argued that there is no syntactic deficit in the L2 syntax, but rather that a deficit exists in morphology. Because communication pressures cannot account for the qualitatively similar pattern in production and comprehension, she proposed that errors reflect the possible insertion of underspecified forms (i.e., default forms that may be inserted instead of a fully specified form under certain conditions in which reducing cost is more important than satisfying feature matching) under a feature-geometric model (e.g., Harley & Ritter, 2002), but that cases of feature clash (i.e., forms that, due to feature mismatch, are ungrammatical) are avoided.

### *2.3. Addressing the debate*

In comparing the results of similar tasks in Franceschina (2005) and McCarthy (2008), the strikingly similar patterns may allow for multiple interpretations: McCarthy's conclusions may allow for another interpretation of Franceschina's (2005) results and vice versa. Indeed, the learners in Franceschina's (2005) study demonstrated accuracy rates above 75%. Yet, she argued that the uninterpretable gender feature is absent in the L2 grammar. And even though the learners in McCarthy's (2008) study only obtained between 60% and 70% accuracy, she argued that the uninterpretable gender feature is part of the L2 grammar. Franceschina rejected an explanation such as the Missing Surface Inflection Hypothesis (Prévost & White, 2000), because this hypothesis proposed that the problem occurred in the mapping from the syntax to the morphology or in the access to the correct forms in real time production as a result of communication pressure, yet she did not find an asymmetry in accuracy rates between production and comprehension tasks. However, the results from the comprehension data in McCarthy's study undermine this argument. Indeed, the computational model argued for by McCarthy that relies on a hierarchy of features along with a dissociation between syntax and morphology can account for Franceschina's findings as well. Unfortunately, these offline studies do not provide evidence on the manner in which learners extract and make use of the relevant features from the input and thus do not allow us to tease apart the two hypotheses examined here, because they differ in their predictions at the processing level. Indeed, the Interpretability Hypothesis (e.g., Tsimpli & Dimitrakopoulou, 2007) predicts asymmetries only between cases of match and mismatch, whereas the Feature Reassembly Hypothesis (e.g., Lardiere, 2009) also predicts an asymmetry between the two cases of mismatch—namely, feature clash and underspecification—based on a hierarchy of features. Therefore, the present study relies on online measures to enable a direct test of the predictions of these two hypotheses, since reading-time (RT)-evidence can provide a window into the nature of the mechanisms used by L2 learners in extracting agreement information in real time.

## 3. Processing studies on L2 gender agreement

To measure learners' sensitivity to a grammatical point, researchers have turned to processing evidence. Many studies have relied on RT data to investigate a variety of phenomena (e.g., Juffs &

Harrington, 1995). In such cases, RTs are gathered as L2 learners complete a reading task, often coupled with a judgment or a comprehension task. Specific asymmetries based on the manipulation of one aspect of the sentence, such as gender agreement, are then examined. For instance, studies that look at gender agreement would expect to find longer RTs for sentences with a gender agreement violation, which reflects the difficulty in processing the incorrect form, compared to sentences with gender agreement. We now turn to two studies that have used such online measures.

Foote (2011) examined the sensitivity to grammatical gender of high proficiency English learners of Spanish in a self-paced moving-window reading task. There were two groups of learners—namely, early and late bilinguals ($n$ = 20 in each group), with the first group classified as heritage speakers—as well as a group of 20 native speakers. The task examined both subject-verb agreement (examining number) and noun-adjective agreement (examining gender), in two conditions: with or without intervening material between the noun and the adjective or the subject and the verb. To ensure that participants were reading the sentences for meaning, half of the sentences were followed by a comprehension question. Results indicated that all groups did comprehend the sentences. For sentences with noun-adjective agreement specifically, two segments were examined: the segment manipulated for (un)grammaticality (i.e., the adjective) and the segment immediately after to check for any spill-over effects. In all participant groups, RTs were significantly longer for the ungrammatical than the grammatical conditions ($p < .001$), no matter if the agreement was separated by intervening materials. Note also that the native speakers and the early bilinguals slowed down on both segments ($p < .05$ and $p < .01$, respectively), whereas the late bilinguals only showed a delayed effect ($p < .01$). Overall, the findings from this study suggest that learners, like native speakers, are sensitive to the uninterpretable gender feature during reading.

In a study that examined the effect of noun animacy in online performance, Sagarra and Herschensohn (2011) collected RTs from 69 beginning and 64 intermediate L1 English-L2 Spanish learners. The learner groups, along with a group of 63 Spanish native speakers, were given a self-paced moving-window reading task paired with yes-no comprehension questions. The results of the comprehension questions revealed a significant main effect for animacy for all three groups ($p < .01$), reflecting the fact that all groups were more successful in responding to questions with inanimate nouns than to those with animate nouns. The RT results indicate a gender congruency and a noun animacy effect only in the intermediate learners and the Spanish native speakers, but not in the beginner learner group. This effect was apparent from the longer RTs on the adjective and the preposition directly following the adjective in sentences with gender violations compared to sentences with gender agreement ($p < .01$). These findings led Sagarra and Herschensohn to suggest that the intermediate learners' grammar relied on language-specific mechanisms involving the uninterpretable gender feature during reading. In terms of the debate on the accessibility of features examined here, these results would thus support an account such as the Feature Reassembly Hypothesis (e.g., Lardiere, 2009), because learners arguably exhibit morphosyntactic knowledge. However, proponents of the Interpretability Hypothesis (e.g., Hawkins & Casillas, 2008; Tsimpli & Dimitrakopoulou, 2007) may argue that these intermediate learners had been exposed to sufficient input to arrive at more automatized context-sensitive rules of feature associations, thus explaining the qualitatively similar patterns of intermediate learners and native speakers. Additionally, that the beginning learners' data did not provide evidence for sensitivity to the gender feature seems to indicate that learners may be confined only to the uninterpretable features present in their L1 grammar.

In sum, although studies like these are important in determining L2 behavior with gender agreement in online studies, they do not specifically examine the debate between the Interpretability Hypothesis (Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007) and the Feature Reassembly Hypothesis (Lardiere, 2009). The current study aimed to provide a test of the predictions made by these hypotheses.

## 4. Research questions and predictions

We have discussed here only two recent studies that have examined gender agreement in L2 Spanish; however, other studies on L2 Spanish and other languages have investigated this phenomenon, using a variety of online techniques (e.g., Keating, 2009; López Prego & Gabriele, 2012;

Tokowicz & MacWhinney, 2005, for Spanish; Dekydtspotter & Renaud, 2009, for French; Hopp, 2012, for German). Overall, these studies have demonstrated that learners, like native speakers, are sensitive to the uninterpretable gender feature. However, only one study has examined the debate of interest here, in L1-English L2-French (Renaud, 2010). The results from that study provided support for the Feature Reassembly Hypothesis. One goal of the present study was to provide an additional language pairing that will lend support to one of the hypotheses, which should hold true regardless of the language examined. Hindi was chosen as the target language of the present study because it exhibits grammatical gender and examining data from English learners of L2 Hindi presents a similar scenario to examining data from English learners of L2 Spanish or English learners of L2 French. Gender agreement in Hindi is accomplished in essentially the same manner as in Spanish: All nouns are assigned either a natural gender or an unrelated but inherent grammatical gender. Adjectives agree with nouns in attributive constructions. Like in Spanish, gender agreement can be regular (or canonical) or irregular. In terms of feature organization, in Hindi, the masculine gender, taken to be the default, does not need to be specified, whereas the feminine gender is marked for gender. Notable differences between Hindi and Spanish are that (a) a non-Roman script is used in writing, and (b) the adjective always precedes the noun (Hindi is a head-final language). Here, only cases of canonical, overt agreement were examined: The masculine noun and adjective ending is आ, whereas the feminine ending is ई, as shown in (1) and (2), respectively. In the plural, only masculine adjectives mark for both gender and plurality, as shown in (3). In contrast, feminine adjectives only exhibit the gender agreement, as the plural is not marked overtly, as in (4).

(1) एक अच्छा लड़का स्कूल जाता है |
A good-MASC boy-MASC school goes
"A good boy goes to school."

(2) एक अच्छी लड़की स्कूल जाती है |
A good-FEM girl-FEM school goes
"A good girl goes to school."

(3) दो अच्छे लड़के स्कूल में हैं|
Two good-MASC-PL boys-MASC-PL school in are
"Two good boys are in school."

(4) दो अच्छी लड़कियाँ स्कूल में हैं|
Two good-FEM-PL girls-FEM-PL school in are
"Two good girls are in school."

The following research questions guided this study: Do L2 Hindi learners of a native language without grammatical gender (i.e., English) exhibit sensitivity to this uninterpretable feature? And more specifically, does the learners' processing profile as measured by RT data support a domain-general (as argued by the Interpretability Hypothesis) or a domain-specific (as argued by the Feature Reassembly Hypothesis) account?

The Interpretability Hypothesis (e.g., Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007) predicts that, because the uninterpretable gender feature is absent from the L1 grammar, learners will rely only on domain-general rules of co-occurrences. It is thus expected that matching features will yield lower costs (i.e., shorter RTs) than mismatching features. Similarly, acceptance rates should reveal a contrast between matching versus mismatching features: The former should yield high acceptance rates, whereas the latter should yield low acceptance rates. Crucially, such a hypothesis does not predict that learners will treat the mismatching forms (i.e., underspecified and clashing forms) differently, because mismatches are absent in the input. Finally, since this hypothesis relies on information gathered from the input, frequency may also play a role.

The Feature Reassembly Hypothesis (e.g., Lardiere, 2009) predicts that learners will rely on UG-constrained operations to compute the gender feature. Moreover, following from the feature-geometric model (e.g., Harley & Ritter, 2002), asymmetries resulting from the specifications of the gender

feature are expected to arise. In other words, the Feature Reassembly Hypothesis will be supported by longer RTs in cases of feature clash (i.e., a feminine form in a masculine context), reflecting a higher processing cost. In contrast, shorter RTs should occur in cases of underspecification (i.e., a masculine form in a feminine context), reflecting fewer costs in the computation. Cases of agreement should yield intermediate RTs, reflecting the cost of Agree. In terms of acceptance rates, it is predicted that learners will accept agreeing cases more than disagreeing cases. However, it is possible that these effects may not yet be visible at lower proficiency levels if the parser guides acquisition.

## 5. A pilot study

### *5.1. Participants*

In the United States, offering Hindi at the collegiate level is fairly rare, as it is a lesser studied language. Many students enrolled in these courses come from an Indian background (i.e., heritage learners), and therefore the population of English L2 learners of Hindi is rather restricted. The participants in the current study were all undergraduate students (except for three native speakers who were graduate students) at a large Southwestern and a large Midwestern university: nine low-intermediate L2 learners and ten native Hindi speakers. It should be noted that all Hindi native speakers were also speakers of English living in the United States, and four of them were also competent in another Indian language. The L2 learners were English native speakers who began taking Hindi courses at the university and were all in their second or third semester of Hindi. However, all had studied a L2 in high school (six Spanish, one German, and one French) for varying amounts of time (from 2 to 8 years). Note also that the participants indicated that they had stopped learning that language by the time they started learning Hindi. After looking at the individual performance on the c-test and on the accuracy in acceptability judgments, years of exposure to another language with gender did not seem to have an effect on either score. For instance, the person who had spent the most time studying another language with gender did not show higher accuracy in either measure than a person who had only studied it for 2 years. Because all learners had a similar exposure to a language with gender, the influence is assumed to be similar across participants. However, it is possible that such an exposure may have enhanced processing effects because learners may have been more attuned to the gender agreement phenomenon.

Participants first completed a short language history questionnaire along with a c-test of intermediate-level material as an independent test of proficiency (Klein-Braley, 1985). The c-test consists of a text in which the second part of certain words has been removed. In this case, the text was taken from a short narrative from the course textbook, *Hindi: A Complete Course for Beginners* (Living Languages, 2007). The endings of 31 critical verbs and nouns that checked for gender and number agreement were removed, and participants were given between 15 and 20 minutes to fill in the ends of words. A point was assigned to each correct response, and zero points were provided for incorrect answers, yielding a maximum of 31 points. On average, the learners scored 19.55 and the native speakers 29.3. Along with participants' self-rating of their Hindi language skills, the c-test results confirmed the proficiency level of the learners to be beginning/low-intermediate.

### *5.2. Materials and Procedure*

The main task was a self-paced moving-window reading task combined with a grammaticality judgment task. The stimulus consisted of seven quadruples in both the singular and the plural conditions, resulting in 56 experimental items (28 per condition). Recall that, in the singular, gender is overtly marked on adjectives in both the masculine and the feminine. In the plural, gender is once again marked on both masculine and feminine adjectives, but only masculine adjectives are also overtly marked for number. Note also that all nouns selected in this study exhibit both canonical masculine and feminine endings as well as distinct singular and plural endings. The experimental items were interspersed within 56 filler items, which were composed of sentences that looked at other

structures, such as comparisons,[1] as well as sentences that focused on semantic correctness. The sentences were intended to be undemanding and were controlled for length and proficiency level (i.e., using introductory vocabulary). The 56 experimental sentences were manipulated for gender (agreement vs. disagreement) on the attributive adjective that immediately preceded the noun. Example quadruple items are provided in Figures 1 and 2 for singular and plural, respectively.

*Figure 1*. Example singular quadruple item

| Context sentence | |
|---|---|
| हम स्कूल के लिए सुबह-सुबह उठे हैं । | |
| We school for morning-morning wake up | |
| "We wake up for school in the early morning." | |
| a. MASC adjective - MASC noun | b. *FEM adjective - MASC noun |
| एक अच्छा लड़का स्कूल जाता है । | एक अच्छी लड़का स्कूल जाता है । |
| A good-MASC boy-MASC school goes | A good-FEM boy-MASC school goes |
| "A good boy goes to school." | "A good boy goes to school." |
| c. FEM adjective - FEM noun | d. ?MASC adjective - FEM noun |
| एक अच्छी लड़की स्कूल जाती है । | एक अच्छा लड़की स्कूल जाती है । |
| A good-FEM girl-FEM school goes | A good-MASC girl-FEM school goes |
| "A good girl goes to school." | "A good girl goes to school." |

*Figure 2*. Example plural quadruple item

| Context sentence | |
|---|---|
| आज दिल्ली में सोमवार की सुबह है। | |
| Today Delhi in Monday morning is | |
| "Today in Delhi, it is Monday morning." | |
| a. PLU adjective - PLU noun | b. *SING adjective - PLU noun |
| दो अच्छे लड़के स्कूल में हैं। | दो अच्छी लड़के स्कूल में हैं। |
| Two good-MASC boys-MASC school in are | Two good-FEM boys-MASC school in are |
| "Two good boys are in school." | "Two good boys are in school." |
| c. SING adjective - PLU noun | d. ?PLU adjective - PLU noun |
| दो अच्छी लड़कियाँ स्कूल में हैं। | दो अच्छे लड़कियाँ स्कूल में हैं। |
| Two good-FEM girls-FEM school in are | Two good-MASC girls-FEM school in are |
| "Two good girls are in school." | "Two good girls are in school." |

In the moving-window judgment task, participants first completed four practice items that contained two grammatical and two ungrammatical examples to familiarize the participants to the types of sentences in the task. The 112 Hindi sentences were pseudo-randomized into four sections, so that only one item per quadruple appeared within one section. Each sentence was preceded by a context sentence in Hindi, which appeared all at once and was related semantically to the target sentence, as in Figures 1 and 2. Following the context, participants read the experimental sentence that appeared word-by-word; however, sometimes two related words appeared together, like compound verbs or prepositional phrases. (In examples [5] and [6], the segmentation of a sentence is indicated with the slashes.) In order to continue reading the sentence and see the next word, participants were asked to accept or reject each word as it appeared. A participant would accept the word by pressing the

[1] An example of a context followed by a comparison sentence is as such: राधा घर में सबसे बड़ी लड़की और सीता सबसे छोटी लड़की है । "At home, Radha is the biggest girl and Sita is the smallest girl." राधा सीता से लम्बी है । "Radha is taller than Sita."

right arrow if he or she found no logical or grammatical error. In contrast, a participant would reject the word on the screen by pressing the left arrow, if he or she found a grammatical error or if the target sentence did not seem to logically follow from the context. In such cases, participants were to continue to label each word as ungrammatical until the end of the target sentence by continuing to press the left arrow. This technique of incorporating a grammaticality judgment task within a self-paced reading task was used to help pinpoint participants' awareness of the location of the problem as they continued to either accept or reject the sentence until the next context sentence appeared on the screen. Although the combination of a grammaticality judgment task with a self-paced moving-window experiment is not novel (see, e.g., Juffs, 2001; Mitchell, 2004), in this particular task, participants were asked not only to judge the sentence once it was read but as they were reading (for each segment). We decided to use this novel technique to see whether we could obtain more fine-grained results (i.e., the exact point at which participants found the sentence to be ungrammatical) rather than a general end-of-sentence measure. To ensure that the participants were not solely focused on the grammaticality judgments, they were to accept or reject English sentences ($n$ = 39) that checked for the comprehension of the Hindi sentences. These English sentences were manipulated in terms of the semantic correctness of their relation to the Hindi item. These sentences were distributed evenly among test items and fillers. An example of a test item with a correct comprehension sentence is provided in (5) and one with an incorrect comprehension sentence is given in (6). Both groups performed well on this task, indicating that they understood the sentences they read.

(5) एक / नई / टोपी / बाज़ार में / है |
The / new-FEM / hat-FEM / market in / is
"The new hat is in the market."
English sentence: *The hat is new*.

(6) एक / लम्बा / राजा / पानी / पीता है |
A / tall-MASC / king-MASC / water / drinks
"A tall king drinks water."
English sentence: *The king drinks tea.*

### *5.3. Analysis*

Acceptance rates were scored as follows. An item was deemed accepted when participants rejected no more than one segment by pressing the left arrow. An item was considered to be rejected when a participant had pressed the left arrow at least twice in a row (out of a possible four) from the critical noun onward. Because often only one segment was rejected in the middle, whereas all other segments were accepted, indicating a potential error in button pressing, we adopted a minimum of two rejections in a row from the point at which we had manipulated the grammaticality of the sentence (i.e., the noun) for it to count as being rejected.

RTs were recorded and examined on the critical noun, which follows the adjective in Hindi, as well as on the segment immediately following the noun—that is, the direct object. These segments were scrutinized for two reasons: First, only attributive adjective agreement cases were investigated here, and because the adjective appears before the noun, agreement is expected to be computed upon encountering the noun segment. Second, delayed effects have been found to occur in L2 acquisition; therefore, a complete analysis should include the spill-over segment—that is, the segment immediately after the noun. The data were coded with SPSS, and extreme RTs (e.g., at two standard deviations from the mean and below 200 milliseconds) were eliminated and replaced with the mean for all participants. Such a procedure is commonly used to eliminate outliers. Note also that replacing the empty cells with the new mean allows keeping the original balanced design without altering the data.

For each condition, repeated measures 2x2x2 ANOVAs were run on acceptance rates and RT data with noun (masculine or feminine) and agreement (corresponding to the manipulation of gender agreement—masculine or feminine—on adjectives) as within-subjects factors and group (learners and natives) as between-subjects factor. Such an analysis was conducted to examine whether there were any effects of agreement as well as of group. The significance level was set at .05 for planned $t$-tests.

These *t*-tests were performed to investigate asymmetries between (a) agreeing and disagreeing forms (in both the masculine and feminine cases), (b) the two types of disagreeing forms (i.e., underspecified versus clash), and (c) the masculine agreeing versus the masculine underspecified forms.

## 6. Results

### *6.1. Singular*

Table 1 provides the mean acceptance rates by condition for each group. To distinguish the mismatched forms more easily, we use the star to indicate clashing forms and the question mark to indicate cases of underspecification. Note first that both groups of participants accepted grammatical sentences at high levels. However, neither group rejected the ungrammatical sentences categorically; instead, they accepted ungrammatical sentences at fairly high rates (between 40 and 60%). An ANOVA revealed a significant interaction of noun and agreement, $F(1, 17) = 9.512$, $p < .01$, which suggests that participants (especially the native speakers) were sensitive to the manipulation of gender agreement. Planned *t*-tests further indicated that the native speakers accepted the agreeing, masculine forms more than the clashing, feminine forms, $t(9) = 2.941$, $p < .05$. This group also accepted the agreeing, feminine forms more than the underspecified, masculine forms, $t(9) = 3.236$, $p < .01$. Such results reflect morphological expectations. In contrast, despite a similar pattern in the learner group, the asymmetries were not found to be significant.

*Table 1*. Acceptance Rates (in %)

| Group | Masc-masc | *Fem-masc | Fem-fem | ?Masc-fem |
|---|---|---|---|---|
| NNS | 74.60 (22.35) | 58.73 (35.31) | 79.37 (24.86) | 63.49 (26.83) |
| NS | 77.14 (31.01) | 45.71 (41.40) | 77.14 (28.73) | 37.13 (43.75) |

*Note*: Standard deviations are in parentheses.

The RT data on the critical noun segment are presented in Table 2. An ANOVA yielded a main effect of agreement, $F(1, 17) = 6.906$, $p < .05$, and a significant interaction of agreement and group, $F(1, 17) = 20.624$, $p < .0001$. Once again, these results point to the sensitivity of participants to the uninterpretable gender feature. The *t*-tests further revealed that the learner group read the agreeing, masculine form faster than the clashing, feminine form, $t(8) = 3.848$, $p < .005$. Additionally, they slowed down when the feminine noun followed a feminine, agreeing adjective compared to a masculine (underspecified) adjective, $t(8) = 2.365$, $p < .05$. Finally, this group also behaved differently on the disagreeing cases: They read the clashing forms more slowly than the underspecified forms, $t(8) = 2.667$, $p < .05$. In short, the learner group seemed to rely on the computation of the gender feature, with RT asymmetries between mismatching and matching forms as well as between clashing and underspecified forms. Several asymmetries were also obtained in the native speaker group: They slowed down on the masculine (underspecified) forms compared to the feminine, agreeing forms, $t(9) = 2.657$, $p < .05$. Moreover, the underspecified forms yielded longer RTs than the clashing forms, $t(9) = 4.023$, $p < .005$. An asymmetry between agreeing and underspecified masculine forms was found to be significant in this group: with faster RTs on the agreeing masculine forms than on the underspecified masculine forms, $t(9) = 2.412$, $p < .05$. Surprisingly, the native speakers did not exhibit a difference in RTs on the nouns following agreeing masculine adjectives and those following the clashing feminine adjectives ($p = .579$). Overall, the native speakers also exhibited RT patterns suggestive of the computation of the gender feature.

*Table 2*. RTs (in ms) on the noun segment by group and condition

| Group | Masc-masc | *Fem-masc | Fem-fem | ?Masc-fem |
|---|---|---|---|---|
| NNS | 1356.82 (453.58) | 1683.43 (593.85) | 1761.21 (586.90) | 1400.55 (490.12) |
| NS | 1131.79 (513.33) | 1194.15 (446.03) | 1177.82 (394.83) | 1423.62 (530.31) |

*Note*: Standard deviations are in parentheses.

Because processing delays have often been observed in L2 acquisition research, RTs were examined on the segment immediately following the critical noun, the direct object (see Table 3). However, an ANOVA run on this segment did not reveal any significant asymmetries. This lack of significant effect was further confirmed by the *t*-tests.

*Table 3*. RTs (in ms) on the spill-over segment by group and condition

| Group | Masc-masc | *Fem-masc | Fem-fem | ?Masc-fem |
|---|---|---|---|---|
| NNS | 1212.96 (329.50) | 1095.27 (417.42) | 1181.39 (424.48) | 1051.64 (222.04) |
| NS | 850.58 (345.85) | 805.02 (338.47) | 879.04 (316.87) | 771.50 (213.80) |

*Note*: Standard deviations are in parentheses.

## *6.2. Plural*

The results from the acceptance rates of the plural items are shown in Table 4. The pattern of high acceptance rates for grammatical sentences in the singular condition holds true for this condition. Once again, both the learners and the native speakers accepted sentences with mismatch at a high rate, around chance-level. As in the singular condition, the ANOVA yielded a significant interaction of noun and agreement, $F(1, 17) = 11.517$, $p < .005$. In the learner group, the planned *t*-tests revealed that the learners accepted the agreeing feminine forms more than the underspecified masculine forms, $t(8) = 3.536$, $p < .01$. However, surprisingly, the difference in acceptance rates between the masculine agreeing forms and the feminine clashing forms was not significant for this group ($p = .171$). The same pattern was obtained in the native speaker group: They accepted more agreeing feminine forms than underspecified masculine forms, $t(9) = 2.529$, $p < .05$, and tended to accept more agreeing masculine forms than clashing feminine forms, $t(9) = 2.193$, $p = .056$.

*Table 4*. Acceptance rates (in %)

| Group | Masc-masc | *Fem-masc | Fem-fem | ?Masc-fem |
|---|---|---|---|---|
| NNS | 77.78 (25.86) | 60.32 (29.26) | 82.54 (17.27) | 58.73 (24.16) |
| NS | 72.86 (39.53) | 48.57 (40.52) | 74.39 (34.21) | 48.57 (42.63) |

*Note*: Standard deviations are in parentheses.

Table 5 presents the mean RTs for the noun segment. The ANOVA yielded a main effect of noun, $F(1, 17) = 27.431$, $p < .0001$, as well as a significant interaction between noun and agreement, $F(1, 17) = 25.587$, $p < .0001$. The planned *t*-tests run on this segment for the learner group indicated that this group read the agreeing, feminine forms faster than the underspecified, masculine forms, $t(8) = 4.836$, $p < .001$. They also read the agreeing, masculine form more quickly than the clashing, feminine form, $t(8) = 3.306$, $p < .05$. The same asymmetries emerged in the native speaker group: with faster RTs on the agreeing, feminine forms compared to the underspecified, masculine form, $t(9) = 3.962$, $p < .005$, and faster RTs on the agreeing, masculine forms compared to the clashing, feminine forms, $t(9) = 4.294$, $p < .005$. In short, both groups exhibited differences in RTs between agreeing and disagreeing forms on this segment.

*Table 5*. RTs (in ms) on the noun segment for both groups

| Group | Masc-masc | *Fem-masc | Fem-fem | ?Masc-fem |
|---|---|---|---|---|
| NNS | 1221.30 (441.25) | 1959.16 (882.55) | 1907.03 (779.71) | 2327.10 (1024.91) |
| NS | 923.91 (435.38) | 1242.85 (472.29) | 1185.37 (328.05) | 1533.24 (837.57) |

*Note*: Standard deviations are in parentheses.

To check for any delayed effect, the segment immediately following the noun was examined (see Table 6). An ANOVA revealed several significant interactions: Noun * Group, $F(1, 17) = 4.772$, $p < .05$, Noun * Agreement, $F(1, 17) = 8.489$, $p < .01$, and Noun * Agreement * Group, $F(1, 17) = 5.750$, $p < .05$. These interactions result from delayed effects obtained in the learner group. Three asymmetries were found to be significant in the *t*-tests for the learners: (a) slower RTs on agreeing, feminine forms

compared to the underspecified masculine form, $t(8) = 4.260$, $p < .005$, (b) faster RTs on underspecified, masculine forms compared to the clashing, feminine forms, $t(8) = 2.343$, $p < .05$, and (c) faster RTs on underspecified, masculine forms compared to agreeing masculine forms, $t(8) = 4.501$, $p < .005$. No significant results were obtained in the native speaker group.

*Table 6*. RTs (in ms) on the spill-over segment

| Group | Masc-masc | *Fem-masc | Fem-fem | ?Masc-fem |
|---|---|---|---|---|
| NNS | 1251.28 (378.98) | 1230.78 (260.29) | 1288.51 (279.85) | 963.08 (272.25) |
| NS | 753.39 (243.56) | 762.59 (217.77) | 804.80 (302.29) | 762.02 (239.04) |

*Note*: Standard deviations are in parentheses.

## 7. Discussion and conclusion

### *7.1. Summary of results*

Overall, the participants were very good at accepting sentences with correct gender agreement (above 75%) but at the same time seemed to lack the ability to reject incorrect sentences in a categorical manner. These acceptance rate patterns are similar in the singular and the plural conditions. The asymmetries between the acceptance of matching and mismatching forms were significant in the native speaker group overall, but only the difference between underspecified masculine and agreeing feminine forms in the plural reached significance in the learner group.

The RT results on both the noun and the segment immediately following it suggest that learners and native speakers were sensitive to the uninterpretable gender feature. Indeed, an interaction with agreement was found in the singular and the plural items. Several asymmetries were in evidence in the learner group: In the singular items, they exhibited asymmetries between agreeing and disagreeing forms and between clashing and underspecified forms. In the plural, the same asymmetries were found in addition to slower RTs on agreeing masculine compared to underspecified masculine forms as a delayed effect. The native speaker group exhibited only one asymmetry in the singular, between underspecified masculine and agreeing feminine forms. In the plural, however, this group revealed faster RTs on agreeing forms than on mismatching forms (both underspecified and clashing forms).

### *7.2. On the two hypotheses*

In response to our first research question, which asked whether learner of a native language without grammatical gender (i.e., English) can exhibit sensitivity to this uninterpretable feature, the results suggest that they do indeed show RT (and acceptance rates) patterns stemming from the manipulation of gender agreement. Such a finding is in line with previous research that has found that learners are sensitive to the uninterpretable gender feature in online tasks (e.g., Dekydtspotter & Renaud, 2009; Foote, 2001; Hopp, 2012; Keating, 2009; López Prego & Gabriele, 2012; Renaud, 2010; Sagarra & Herschensohn, 2011; Tokowicz & MacWhinney, 2005). The second question aimed at teasing apart the claims of the two hypotheses—namely, the Interpretability Hypothesis (e.g., Tsimpli & Dimitrakopoulou, 2007) and the Feature Reassembly Hypothesis (e.g., Lardiere, 2009). More specifically, we asked whether the learners' processing profile as measured by RT data supports a domain-general or a domain-specific account.

Both the Interpretability Hypothesis and the Feature Reassembly Hypothesis predict that learners will accept agreeing cases more than disagreeing cases. This result was generally borne out in the data and is consonant with previous research (e.g., Franceschina, 2005; McCarthy, 2008). However, other findings were a bit surprising and deserve mention here. First, learners and native speakers alike accepted the mismatching forms at fairly high rates overall (up to 60%). Such a finding was unexpected for the native speakers, as they should have rejected such forms almost categorically. It is certainly unclear as to why the results are so high. It could be an effect of pleasing the researcher, by agreeing to all items presented. Another possibility is that the native speakers we were able to recruit were not necessarily used to reading the Hindi script, and it may have led to their lack of determinate judgments. Note that such an explanation may also account for the slow RTs and the rather flat RT

patterns observed. As for the learners, who had a low-intermediate level of proficiency, the fairly high acceptance rates of the mismatched forms may simply reflect a delay in their grammatical knowledge. A similar delay in grammar compared to processing behavior has been obtained in intermediate learners of French (see Renaud, 2010) and in learners of Spanish (see Tokowicz & MacWhinney, 2005). Crucially, the asymmetries in acceptance rates are in the direction predicted by both hypotheses: higher acceptance of matching forms. An anonymous reviewer pointed out that participants may have relied on a strategic interpretation. It is unclear to us how this could have occurred since the items presented to the participants were interspersed among an equal number of filler items of different structures, including items that were grammatically correct but semantically impossible and vice versa. Additionally, because participants were unaware of the purpose of the experiment, it is highly unlikely that they were able to rely on a strategy to respond to the gender manipulation.

As for the processing results, recall that both the Interpretability Hypothesis (e.g., Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007) and the Feature Reassembly Hypothesis (Lardiere, 2009) predict differences in the way learners treat ungrammatical and grammatical sentences. This prediction was also borne out. Indeed, RT asymmetries between matching and mismatching forms were obtained, albeit not all such asymmetries reached significance. Most surprising is the fact that the native speakers did not exhibit a significant asymmetry between agreeing masculine and clashing feminine forms with the singular items. The same surprising finding in the learner group may stem from their lack of sensitivity. However, because the natives exhibited the same pattern, an alternative explanation would need to be considered. We suggest that this lack of effect may be due to the rather slow RTs, which may have masked the distinction between the two forms. To distinguish between the two proposals, it is necessary to examine the participants' online performance on the two types of mismatch as well as on agreeing and underspecified masculine forms.

The Interpretability Hypothesis (e.g., Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007) does not predict a difference in performance between the underspecified and clashing forms. As Hawkins and Casillas (2008) have recently proposed, when faced with an uninterpretable feature not licensed by the L1 grammar (such as the gender feature in this study), learners will rely on context-sensitive clues and distributional regularities in the input. Learner input consists of grammatical sentences, from which the learners parse relevant information that triggers a grammatical change in their interlanguage. In other words, frequency of occurrence should also provide more opportunities of the creation of rules of co-occurrence. This frequency effect would presumably translate into faster RTs for matching cases, suggestive of such automatized feature-associative routines. However, because ungrammatical sentences are absent from the input, they should treat both types of ungrammaticality (clashing and underspecified forms) similarly. Even though the acceptance rates seem to provide support for such an account, the RT data suggest that the learners treated the underspecified and the clashing forms differently, at least in the singular condition, in which they read the clashing forms more slowly than the underspecified forms. The Feature Reassembly Hypothesis (Lardiere, 2009), in contrast, expects an asymmetry between the underspecified and clashing forms as a result of the hierarchy of features (Harley & Ritter, 2002). Indeed, underspecified forms should create the fewest costs, whereas clashing forms would result in the highest processing costs reflective of the clash in the derivation. This scenario corresponds to the findings in the present study. Additionally, the fact that such effects were obtained in the processing of the learners, whereas their acceptance rates seemed at chance on these conditions, points to the role of the parser in guiding L2 acquisition. Note that these findings parallel those obtained in L2 French (see Renaud, 2010).

Another asymmetry that is difficult to explain for the Interpretability Hypothesis is the different RTs between the agreeing and the underspecified masculine forms. Indeed, if one assumes that underspecified masculine forms are merely mismatching forms, then these forms should be read more quickly overall. However, both the learners and the native speakers showed the opposite pattern. Moreover, if one assumes that the acquisition of the forms is driven primarily by frequency information, because these forms are all masculine forms, they should be treated similarly. Once again, this is not what was found. The Feature Reassembly Hypothesis in contrast would predict a difference between both forms as a result of the hierarchy of features. Problematic for this hypothesis is the direction of the asymmetries. However, it could be argued that the underspecified forms triggered

longer RTs because the forms were considered as possible forms to be inserted in the computation, whereas the agreeing masculine forms triggered a rule of insertion of the masculine value necessary to compute agreement upon encountering the matching feature on the noun (e.g., Harley & Ritter, 2002).

### *7.3. Limitations and implications*

Several limitations need to be mentioned here. First, it should be pointed out that more significant results may have been found had the sample size be larger. However, the findings of this pilot study are nonetheless informative and provide the beginning of an answer to our research questions. A second limitation of the present study is the fact that, to our surprise, some of the participants, especially the native speakers, were not very familiar with the Hindi script. This showed not only in the slow RTs overall but this may also have masked important asymmetries in RTs, as participants may have read each segment at a slower pace. Finally, the methodology used in this study may also have affected some of the asymmetries because participants were asked to judge each segment as they read them. Such a metalinguistic task may have caused for effects to be lost along the way, even though the goal of using such a methodology was to try to find out at which point participants reject sentences in which gender agreement was manipulated. Effects of a secondary task on RT patterns have recently begun to be investigated (see Leeser, Brandl, & Weissglass, 2011).

Despite these limitations, the results from this study seem to confirm that learners are sensitive to a feature that is absent in their L1. Moreover, the Feature Reassembly Hypothesis (Lardiere, 2009) appears to provide the best account for the pattern of response obtained. The processing profile of the L2 learners in this study further suggests that interlanguage grammars, like L1 grammars, rely on the domain-specific checking of features, even though this feature may be absent from the L1 grammar (see also Dekydtspotter & Renaud, 2009). Of course, the results from the present study are tentative given the small sample size; however, these findings provide additional support for the fact that L2 grammars are not different from L1 grammars. Further studies that focus on discovering the mechanisms at play during processing of agreement relations by L2 learners will be able to add to the findings of the present study.

## References


Alarcón, Irma V. (2009). The processing of gender agreement in L1 and L2 Spanish: Evidence from reaction time data. *Hispania*, *92*(4), 814-828.

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

Dekydtspotter, Laurent. (2001). The universal parser and interlanguage: Domain-specific mental organization in the comprehension of combien interrogatives in English-French interlanguage. *Second Language Research, 17*(2), 91-143.

Dekydtspotter, Laurent, & Renaud, Claire. (2009). On the contrastive analysis of features in second language acquisition: Uninterpretable gender on past participles in English-French processing. *Second Language Research, 25*(2), 251-263.

Fodor, Janet D. (1998). Parsing to learn. *Journal of Psycholinguistic Research, 27*(2), 339-374.

Fodor, Jerry A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Foote, Rebecca. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics, 32*(1), 187-220.

Franceschina, Florencia. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. Amsterdam: Benjamins.

Halle, Morris, & Marantz, Alec. (1993). Distributed Morphology and the Pieces of Inflection. In Kenneth Hale & S. Jay Keyser (Eds.), *The View from Building 20* (pp. 111-176). Cambridge, MA: MIT Press.

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

Hawkins, Roger, & Casillas, Gabriela. (2008). Explaining frequency of verb morphology in early L2 speech. *Lingua, 118*(4), 595-612.

Hawkins, Roger, & Chan, Cecilia Yuet-hung. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis.' *Second Language Research, 13*(3), 187-226.

Hawkins, Roger, & Liszka, Sarah. (2003). Locating the source of defective past tense marking in advanced L2 English speakers. In Roeland van Hout, Aafke Hulk, Folkert Kuiken, & Richard Towell (Eds.), *The interface between syntax and lexicon in second language acquisition* (pp. 21-44). Amsterdam: Benjamins.

Hazdenar, Belma, & Schwartz, Bonnie D. (1997). Are there optional infinitives in child L2 acquisition? In Elizabeth Hughes, Mary Hughes, & Annabel Greenhill (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development* (pp. 257-268). Somerville, MA: Cascadilla Press.

Hopp, Holger. (2012). The on-line integration of inflection in L2 processing: Predictive processing of German Gender. In Alia Biller, Esther Y. Chung, & Amelia E. Kimball (Eds.), *Proceedings of the 36th Boston University Conference on Language Development* (pp. 226-245). Somerville, MA: Cascadilla Press.

Juffs, Alan. (2001). Psycholinguistically-oriented second language research. *Annual Review of Applied Linguistics, 21*(21), 207-223.

Juffs, Alan, & Harrington, Michael. (1995). Parsing effects in L2 sentence processing: Subject and object asymmetries in *wh*-extraction. *Studies in Second Language Acquisition, 17*(4), 483-516.

Keating, Gregory D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*(3), 503-535.

Klein-Braley, Christine. (1985). A cloze-up on the c-test: A study in the construct validation of authentic tests. *Language Testing, 2*(1), 76-104.

Lardiere, Donna. (2008). Feature assembly in second language acquisition. In Juana Liceras, Helmut Zobl, & Helen Goodluck (Eds.), *The role of features in second language acquisition* (pp. 106-140). Mahwah, NJ: Erlbaum.

Lardiere, Donna. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research, 25*(2), 173-227.

Leeser, Michael J., Brandl, Anel, & Weissglass Christine. (2011). The role of task and structure in L2 sentence processing. In Kim McDonough & Pavel Trofimovich (Eds.) *Insights from psycholinguistics: Applying priming research to L2 learning and teaching* (pp. 179-198). Amsterdam: Benjamins.

Living Languages. (2007). *Hindi: A complete course for beginners*. New York: Random House.

López Prego, Beatriz, & Gabriele, Alison. (2012). Examining the nature of variability in gender and number agreement in native and non-native Spanish. In Alia Biller, Esther Y. Chung, & Amelia E. Kimball (Eds.), *Proceedings of the 36th Boston University Conference on Language Development* (pp. 315-327). Somerville, MA: Cascadilla Press.

McCarthy, Corinne L. (2007). *Morphological variability in second language Spanish*. Unpublished doctoral dissertation, McGill University, Montréal.

McCarthy, Corinne L. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research, 24*(4), 459-486.

Mitchell, Don C. (2004). On-line methods in language processing: Introduction and historical review. In Manuel Carreiras & Charles Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond.* (pp. 15-32). New York: Psychology Press.

Prévost, Philippe, & White, Lydia. (2000). Missing surface inflection hypothesis or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16*(2), 103-133.

Renaud, Claire. (2010). *On the nature of agreement in English-French acquisition: A processing investigation in the verbal and nominal domains*. Unpublished doctoral dissertation, Indiana University.

Sagarra, Nuria, & Herschensohn, Julia. (2011). Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning, 61*(1), 80-116.

Tokowicz, Natasha, & MacWhinney, Brian. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition, 27*(2), 173-204.

Tsimpli, Ianthi Maria, & Dimitrakopoulou, Maria. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research, 23*(2), 215-242.

White, Lydia, Valenzuela, Elena, Kozlowska-Macgregor, Martyna, & Leung, Yan-kit Ingrid. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics, 25*(1), 105-133.

# Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges between Disciplines

edited by
Ryan T. Miller, Katherine I. Martin,
Chelsea M. Eddington, Ashlie Henery,
Nausica Marcos Miguel, Alison M. Tseng,
Alba Tuninetti, and Daniel Walter

**Cascadilla Proceedings Project  Somerville, MA  2014**



ISBN 978-1-57473-464-5 library binding



## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

## Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Renaud, Claire and Lauren Covey. 2014. Sensitivity to Gender Agreement in Second Language Hindi: A Processing Investigation. In *Selected Proceedings of the 2012 Second Language Research Forum*, ed. Ryan T. Miller et al., 51-64. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3085.