

FEATURES AND FORMS IN L2 SPANISH VERBAL INFLECTION*

María Cristina Cuervo, Samuel Jambrović, and Ohanna Severo
University of Toronto

Second-language (L2) learners of Spanish often produce errors in verbal morphology. This corpus study explores whether L2 speakers tend to substitute default feature values for marked feature values, such as third person for first person and singular for plural. In a novel contribution to the acquisition literature, we consider whether default-for-marked feature value substitutions are modulated by the morphophonological form of third-person singular agreement, which is overt in some tenses (*habl-ó.3SG* ‘s/he spoke’) but null in others (*habla-∅.3SG* ‘s/he speaks’). Overall, our findings suggest that L2 inflectional errors stem from within the morphological component of the grammar and display a sensitivity to featural composition but not necessarily to morphophonological exponence.

1. Introduction

The rich inflectional morphology of the Spanish verb encodes several contrasts through suffixes, including inflectional class, finiteness, mood, tense, person, and number. In their acquisition journey, second-language (L2) learners of Spanish may produce errors related to one or more of these features, as illustrated by the following examples from the Spanish Learner Language Oral Corpora (2020).

- (1) Finiteness (Participant I69SMP13)
*La gente de Inglaterra todavía **hablar** sobre esa princesa.
the people of England still speak.INF about that princess
- (2) Tense (Participant I54NTV13)
*El fin de semana **pasado** yo **hago** el atletismo.
the end of week past I do.PRES the track.and.field
- (3) Person (Participant I65NTV13)
***Yo y mis amigas** **fueron** a la centro de ciudad.
I and my friends went.3PL to the centre of city
- (4) Number (Participant I53SMP13)
*Muchos **niños** **estudia** en el colegio ahora.
Many students study.3SG in the high.school now

*We thank Sophie Harrington for her contributions to all aspects of this study, from coding and formulating our research questions to preparing for the conference. We are also grateful to the members of the University of Toronto’s Syntax Project and the attendees of CLA for their valuable suggestions. This project was partially funded through the Work Study Program of Ontario and a research grant from Victoria College.

In (1), the participant uses an infinitive (*hablar* ‘to speak’) in place of a conjugated verb (*hablo* ‘I speak’). In (2), the participant substitutes the present tense (*hago* ‘I do’) for the preterite (*hice* ‘I did’). In (3), the participant uses third-person agreement (*fuleron* ‘they went’) rather than first-person (*fuimos* ‘we went’), and in (4), singular agreement (*estudia* ‘s/he studies’) rather than plural (*estudian* ‘they study’).

Non-target L2 inflectional production raises two key questions. First, is there a systematicity to errors such that linguistic generalizations can be drawn, or should errors be attributed to performance variability alone? Second, if errors are systematic, do they predominantly stem from a particular component of the interlanguage grammar, such as the syntax, morphology, or phonology?

The literature on Spanish L2 acquisition often considers the kinds of errors that learners produce but not necessarily their source. This work builds on a line of investigation that characterizes inflectional errors as primarily morphological in nature, which is reviewed in section 2. Section 3 introduces our corpus study, and overall findings are presented and discussed in section 4. Section 5 details two case studies that explore the relative importance of morphological features and morphophonological forms with respect to non-target L2 production. Finally, section 6 concludes.

2. Background

Much L2 research focuses on error production as a means of better understanding interlanguage grammars (Lardiere 1998a,b; Prévost and White 2000; Herschensohn 2001; Bruhn de Garavito 2003a,b; McCarthy 2006; Lardiere 2009; McCarthy 2012; Domínguez et al. 2013). In one set of proposals, L2 grammars are argued to be syntactically impaired (Vainikka and Young-Scholten 1994, 1996; Meisel 1997). In another, L2 inflectional errors do not reflect syntactic deficits but are claimed to instead stem from learners’ difficulties mapping abstract morphosyntactic features to their morphophonological forms (Lardiere 1998a,b; Prévost and White 2000; Bruhn de Garavito 2003a,b).

Regarding Spanish, Bruhn de Garavito (2003b) suggests that learners overproduce verbal stems, as in *habla-* ‘speak’, which consist of a root (*habl-*) and a theme vowel (*-a*). Verbal stems have the same phonological form as third-person singular verbs in the present tense (5c), where agreement is null.

- | | | |
|-----|--------------------|----------------------|
| (5) | a. <i>habl-o</i> | d. <i>habl-a-mos</i> |
| | b. <i>habl-a-s</i> | e. <i>habl-á-is</i> |
| | c. <i>habl-a-∅</i> | f. <i>habl-a-n</i> |

In an oral production task, 22 beginner-level L2 Spanish speakers (L1 English) generated 1,324 tokens, 89 (7%) of which contained agreement errors. Of these, 59 (66%) were third-person singular forms—which Bruhn de Garavito analyzes as simplified, uninflected verbal stems rather than exponents of third-person agreement—used in first-person and second-person singular contexts. At least in the present tense, it is possible that morphophonolog-

ical form rather than underlying morphosyntactic features could play a role in L2 errors, a question that has not been fully addressed in the literature.

A different body of work argues that L2 inflectional errors result from the overgeneralization of default feature values (McCarthy 2006, 2012). According to McCarthy's Morphological Underspecification Hypothesis, learners substitute underspecified feature values for specified values (6), not other specified values (7).

(6) Underspecification
 *yo habla-∅
 I speak-3SG

(7) Feature clash
 *yo habla-s
 I speak-2SG

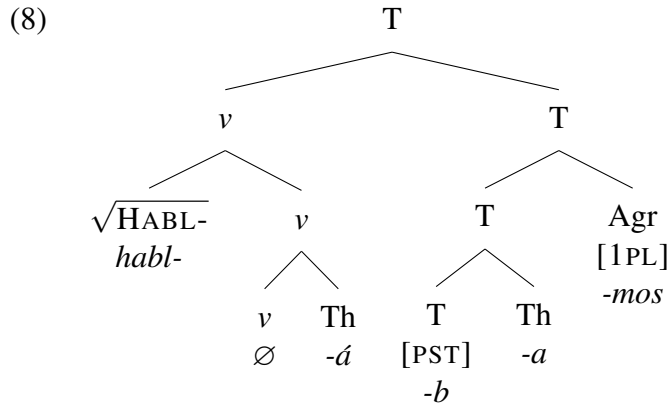
This distinction is based on Harley and Ritter (2002) and Cowper's (2005) proposal that third person is the default value for person, which is the absence of a feature value in a privative system: []. Since *habla* in (6) has an underspecified person value, it does not clash with the specified [1] feature of *yo*, unlike *hablas* in (7), which has its own specified [2] feature. If feature values themselves are the source of errors, the forms that realize these values may be irrelevant, but no study has explicitly made this prediction.

McCarthy (2012) finds evidence that L2 speakers substitute underspecified feature values for specified values along the dimensions of finiteness, person, and number (for which the default values are nonfinite, third person, and singular). Among 960 tokens produced by 37 intermediate-level L2 Spanish speakers (L1 English) in the Spanish Learner Language Oral Corpora, 83 (75%) out of 111 person errors are third-for-first person substitutions, 18 (86%) out of 21 number errors are singular-for-plural substitutions, and 47 (98%) out of 48 finiteness errors are nonfinite-for-finite substitutions.

Our corpus study elaborates on Bruhn de Garavito (2003b) and McCarthy (2006, 2012) in two ways. First, we examine the role of defaults in errors that involve tense as well as those involving finiteness and agreement. Second, we explore whether the overproduction of defaults is affected by morphophonological form, focusing on the case of null versus overt exponents.

2.1 Theoretical assumptions

In the Distributed Morphology framework, vocabulary items are inserted into terminal nodes after syntactic and morphological operations have taken place (Halle and Marantz 1993, 1994; Marantz 1997; Harley and Noyer 1999). The terminal nodes of syntactic structures consist solely of abstract features or feature bundles, such as [PST] and [1PL]. Within the morphological component, and before vocabulary insertion, terminal nodes and their features are subject to operations that include deletion, fusion, morphological merger, and linearization. At the moment of lexical insertion, morphophonological exponents (vocabulary items) compete for insertion into the terminal nodes. We illustrate this process using *hablábamos* 'we were speaking' and Oltra-Massuet and Arregi's (2005: 50) morphological structure of the Spanish verb.



In the syntax, v moves to T to form a complex head (Matushansky 2006). Then, in the morphological component, an agreement position (Agr) is adjoined to T, and a theme position (Th) is adjoined to all functional heads (Oltra-Massuet and Arregi 2005).

Returning to the notion of privative features, the complete paradigm of Spanish verbal agreement morphology can be captured using three features: [1], [2], and [PL].¹

- (9)
- | | |
|-------------------------------|-------------------------------|
| a. [1] first-person singular | d. [1PL] first-person plural |
| b. [2] second-person singular | e. [2PL] second-person plural |
| c. [] third-person singular | f. [PL] third-person plural |

Recall that third person and singular are default, or underspecified, feature values. In Harley and Ritter's (2002) feature geometry, third person is the absence of the participant node, and singular number is the default value for the individuation node.

Lexical insertion follows Halle's (1997) subset principle: the item that is inserted is the one that has the largest subset of features found in the terminal node. The lexical items pertaining to the imperfect tense in Spanish are in (10).

- (10)
- | |
|-------------------|
| a. /-mos/ ↔ [1PL] |
| b. /-is/ ↔ [2PL] |
| c. /-s/ ↔ [2] |
| d. /-n/ ↔ [PL] |
| e. ∅ ↔ [] |

In this tense, if the Agr node contains the feature [1], the null exponent is inserted since the other lexical items all correspond to features that are not a subset of [1], while the empty set is a subset of all sets. As a result, there is syncretism between first-person and third-person singular forms in the imperfect, as in *yo cantaba-∅* 'I was singing' and *él/ella cantaba-∅* 's/he was singing'.

¹The features [1] and [2] are used for simplicity; the same distinctions could be made using features like [PARTICIPANT], [SPEAKER], and [ADDRESSEE].

3. The study

Previous accounts that attribute inflectional errors to the morphological component of interlanguage grammars do not differentiate between abstract feature values and their corresponding exponents in a principled way. Beyond stating that the morphology is compromised, these accounts remain agnostic about the specific source of errors. The purpose of our corpus study is to address this gap in the acquisition literature.

In section 4, we analyze error rates with respect to tense, person, and number to determine whether L2 speakers overgeneralize default feature values. In section 5, we explore whether morphophonological form affects the substitution of defaults by comparing substitution rates for null and overt exponents that spell out default feature values.

3.1 Data

Our data come from the guided interview task of the Spanish Learner Language Oral Corpora (2020). In the first part of this task, participants answered questions about their knowledge of famous historical figures (e.g., William Shakespeare, John Lennon, Diana Spencer), and in the second part, they answered questions about past events in their lives. We analyzed the data from 20 participants who were intermediate-level L2 Spanish speakers (L1 English) in Year 13 of the British educational system, who are typically 17 to 18 years old.

4. Results and discussion

We coded a total of 2,193 verbal tokens from the 20 participants, of which 852 (39%) were errors. All utterances of a verb were counted as separate tokens, including truncations and repetitions. Our rationale for doing so is that participants sometimes produced a correct form followed by an incorrect form, as in (11), just as they often produced an incorrect form followed by a correct form.

- (11) Truncations and repetitions (Participant I56SMP13)
- | | | | | | | | |
|---------|--------------|-----------------|------|--------------|--------|--------------|-------|
| *Cuando | tenía | cuando | tu- | tuve | cuando | tuve | nueve |
| when | had.IPFV.1SG | when | have | had.PRET.1SG | when | had.PRET.1SG | nine |
| años | yo fui | a Florida. | | | | | |
| years | I | went to Florida | | | | | |

The inclusion of repetitions and self-corrections creates a more naturalistic data set, one in which the potential effects of metalinguistic knowledge are mitigated.

Table 1 on the next page lists the number of errors and corresponding percentage for each feature related to verbal inflection; the reason the columns do not add up to 852 and 100% is that many of the non-target produced forms contained errors for more than one feature. To clarify the difference between tense and finiteness, tense refers to cases where a conjugated verb was produced but not in the appropriate tense, while finiteness refers to cases where a nonfinite form (e.g., an infinitive, gerund, or participle) was produced in a finite context, and vice versa.

Table 1. Overall results by error type

Feature	Errors	Percentage of total errors ($N = 852$)
Tense	391	46%
Person	130	15%
Number	97	11%
Finiteness	36	4%
Mood	29	3%
Other	268	31%

Since the remainder of this section addresses tense, person, and number, we start with mood and the “other” category. For mood, there were 2,007 finite targets, with 1,991 (99%) indicative targets and 16 (<1%) subjunctive targets. The correct mood was produced for 1,972 (99%) of the 1,991 indicative targets and for 3 (19%) of the 16 subjunctive targets. Next, errors classified as “other” include truncations, incorrect uses of the copular verbs *ser* and *estar*, and altered or nonexistent roots (e.g., **escribiríamos* rather than *escribiríamos* ‘we would write’), among other issues.

To determine whether L2 errors are systematic, we now examine tense, person, and number in detail, focusing on produced forms that correspond to the default values for each of these features.

4.1 Tense

L2 errors in tense are frequently discussed in studies that focus on the syntax or semantics of tense, and some work has also investigated the role of defaults in L2 errors, either in terms of features or forms. For instance, McCarthy (2012) finds that nonfinite-for-finite substitutions are rare and finite-for-nonfinite substitutions even more so; when considering these values separately, L2 speakers are more accurate with nonfinite morphology than they are with finite morphology. What has yet to be assessed is the role that defaults may play when one finite verb is substituted for another. Just as finiteness is thought to have marked (finite) and unmarked (nonfinite) feature values, so is tense, with present usually argued to be the unmarked value (Cowper 2005, Oltra-Massuet and Arregi 2005).

Of the 2,193 produced forms in our data set, 1,950 (89%) had a target tense of present, preterite, or imperfect. There were an additional 57 (3%) finite targets (future, conditional, and perfect tenses) and 186 (8%) nonfinite targets (infinitives, gerunds, and participles). Example (12) shows an infinitive (*entrar* ‘to enter’) in place of a finite verb (*entran* ‘they enter’), and (13) shows the present tense (*disfruto* ‘I enjoy’) used in a preterite context (*disfruté* ‘I enjoyed’).

- (12) Infinitival-for-finite substitution (Participant I69SMP13)
 *Siempre hay personas que **entrar**.
 always there.are people who enter.INF

(13) Present-for-preterite substitution (Participant I50SMP13)

*Fue un tiempo que que disfruto mucho.
it.was a time that that enjoy.PRES.1SG a.lot

Table 2 contains the number and percentage of errors for the 2,007 finite targets as well as the number and percentages of finite-for-finite and infinitival-for-finite substitutions.

Table 2. Errors involving finite forms

Tense	Targets	Errors	Finite form produced	Infinitive produced
Present	669	32 (5%)	9 (28%)	6 (19%)
Preterite	759	241 (32%)	165 (68%)	8 (3%)
Imperfect	522	244 (47%)	204 (84%)	4 (2%)
Other	57	18 (32%)	14 (78%)	4 (22%)
Totals	2,007	535 (27%)	392 (73%)	22 (4%)

The rate of non-target tense productions reached 27% overall, ranging from 5% for present targets to 47% for imperfect targets. For all tenses except the present, most errors consisted of finite-for-finite substitutions. The rate of infinitival-for-finite substitutions was also unevenly distributed: it was high for present tense targets (19%) and low for preterite (3%) and imperfect (2%) targets. These results challenge the view that the infinitive is the default verbal form in Spanish, which is predicted by the Missing Surface Inflection Hypothesis (Prévost and White 2000).

Next, we take a closer look at the preterite and imperfect errors involving a different finite form to explore whether present is the default value for tense. Table 3 shows that present was indeed the most frequent substitution for the simple past tenses.

Table 3. Tense-related substitutions

Tense	Errors	Present produced	Preterite produced	Imperfect produced
Preterite	165	97 (59%)	–	59 (36%)
Imperfect	204	109 (53%)	89 (44%)	–

The second most frequent type of substitution was one simple past tense for the other. The higher rate of preterite-for-imperfect substitutions is consistent with general findings in the literature that L2 speakers acquire the imperfect later than the preterite and that the imperfect is affected in heritage grammars (Montrul 2011; Domínguez et al. 2013; Montrul 2021; Perpiñán 2021). Furthermore, these figures align with Cuervo and Klassen's (2014) proposal that the imperfect is featurally more specified than the preterite.

4.2 Person

To evaluate person, we coded the 2,007 finite targets in our data set as first, second, or third person. There were 756 (38%) first-person, 3 (<1%) second-person, and 1,248 (62%)

third-person targets.² The example in (14) illustrates two person errors, where a third-person plural verb (*jugaban* ‘they were playing’) and a third-person singular verb (*jugaba* ‘s/he was playing’) are both substituted for a first-person plural verb (*jugábamos* ‘we were playing’).

(14) Third-for-first person substitution (Participant I53SMP13)

*Yo y mi hermano jugaban jugaba en la playa.
I and my brother played.3PL played.3SG on the beach

In total, there were 130 (6%) errors involving person among the 2,007 finite targets. Table 4 indicates that person substitutions were more frequent for first-person targets (13%) than for third-person targets (3%). Assuming that third person is the default feature value and first person is marked, these results show that underspecification errors are more common for person.

Table 4. Errors involving person

Person	Targets ($N = 2,007$)	Errors
First person	756 (38%)	97 (13%)
Second person	3 (<1%)	0 (0%)
Third person	1,248 (62%)	33 (3%)

Of the 97 errors for first-person targets, 94 (97%) were third-person substitutions, and 3 (3%) were second-person substitutions. Of the 33 errors for third-person targets, 30 (91%) were first-person substitutions, and 3 (9%) were second-person substitutions. Among all 130 errors related to person, 94 (72%) were substitutions of the default for a marked feature value, 33 (25%) were substitutions of a marked value for the default, and 3 (2%) were instances of feature clash.

Due to the nature of the guided interview task, there are essentially only two values for person in this data set, and errors for third-person targets inevitably consist of first-person substitutions. As a result, the error percentages in Table 3 only provide part of the picture, and a more informative way of evaluating whether L2 speakers tend to substitute the default value for person is to calculate the proportion of substitutions for first-person and third-person targets relative to the total number of non-target productions (i.e., nonfinite forms, truncations, and other tokens that do not express person). For first-person targets, there were 163 errors overall, including 95 (58%) third-person substitutions and 65 (40%) without a value for person. For third-person targets, there were 122 errors overall, including 34 (28%) first-person substitutions and 85 (70%) without a value for person. The crucial observation here is that person-related substitutions are more frequent than non-person-related substitutions for first-person targets, meaning that L2 errors are sensitive to the featural composition of the target, namely whether it contains marked or default values.

²There were no errors for the second-person targets, so they are excluded from further analysis.

4.3 Number

In our analysis of number, we coded the 2,007 finite targets as singular or plural. There were 1,720 (86%) singular and 287 (14%) plural targets. In (15), the speaker produces a number error by using a third-person singular verb (*ayudó* ‘s/he helped’) rather than a third-person plural verb (*ayudaron* ‘they helped’).

- (15) Singular-for-plural substitution (Participant I57SMP13)
 *Mi padres y mi madre me ayudó olvidar.
 my parents and my mother me helped.3SG to.forget.it

There were a total of 96 (5%) errors involving number among the 2,007 finite targets, which were either singular-for-plural or plural-for-singular substitutions. Table 5 shows that singular-for-plural substitutions (23%) were more common than plural-for-singular substitutions (2%). Based on the view that singular is the default value for number and plural is marked, these results indicate that underspecification is the more common error type for number, just as it is for person.

Table 5. Errors involving number

	Targets ($N = 2,007$)	Errors
Singular	1,720 (86%)	33 (2%)
Plural	287 (14%)	66 (23%)

Since there are only two possible values for number, the error percentages in Table 5 should be supplemented with the proportion of substitutions for singular and plural targets relative to the total number of non-target produced forms (including those that do not express number). For singular targets, there were 33 (22%) plural substitutions among 150 overall errors, and for plural targets, there were 66 (65%) singular substitutions among 102 overall errors. Therefore, number-related substitutions for plural targets are more common than non-number-related substitutions. The analysis of number, like that of person, suggests that L2 errors are sensitive to the featural composition of the target in terms of marked versus default values.

4.4 Key findings

Our results identify a general trend in L2 verbal inflection of overproducing forms that correspond to default feature values, which we observe for finiteness, tense, person, and number. This error pattern is evidenced by higher rates of substitutions using default, or underspecified, feature values, confirming McCarthy’s (2006, 2012) proposal that morphological errors are systematic and that underspecified feature values are substituted more frequently than specified values.

Our analysis of tense-related errors in section 4.1 contributes data for a previously unexplored featural category and enables a consideration of finiteness through the lens of

tense. Among errors in tense, the rate of infinitival-for-finite substitutions was 4% (22 of the 535 finite productions that had non-target tense, as calculated in Table 2), while that of finite-for-infinitival substitutions was 29% (5 of the 17 errors with infinitival targets). These percentages run counter to the proposal that default feature values are most frequently substituted since nonfinite is typically argued to be the default value for finiteness. However, if we factor in tense, the picture changes. Infinitival-for-finite substitutions remain low for targets with a marked value for tense—3% for preterite and 2% for imperfect—but reach 19% for present targets, the default value. For targets that have a marked value for tense, present tense forms account for the highest rate of tense-related errors, not infinitives, including 40% of preterite errors (97 out of 241) and 45% of imperfect errors (109 out of 244). In sum, the type of default feature value that is overproduced (nonfinite for finiteness and present for tense) varies according to value of the target and whether it is marked.

In sections 4.2 and 4.3, we found that default feature values were most frequently substituted for person and number errors as well. Moreover, we determined that the feature value of the target—marked or default—affected the rate of substitutions for that feature: substitutions were more frequent for marked feature values (first person and plural) than for default values (third person and singular). These differences became clear by examining a wider range of errors for person and number, such as nonfinite forms and truncations.

The results in this section clarify two issues. First, they bolster McCarthy's (2006, 2012) argument that L2 inflectional errors most commonly involve the substitution of underspecified feature values for specified values, and they indicate that the type of overused default depends on the featural composition of the target. Second, they challenge Bruhn de Garavito's (2003b) suggestion that learners of Spanish overgeneralize verbal stems given that the default forms of the features under investigation (finiteness, tense, person, and number) do not always match the third-person singular form of the present tense. We now present two case studies that specifically address the role of morphophonological exponence in L2 errors.

5. Case studies

The purpose of the two case studies in this section is to explore whether null exponents that correspond to default feature values are substituted more often than overt exponents. In other words, is the L2 tendency to substitute default feature values—the principal finding of the previous section—modulated by the morphophonological forms that realize these values? Based on data from beginner-level speakers, Bruhn de Garavito (2003b) proposes that learners tend to substitute verbal stems, which are morphophonologically simple (root plus theme vowel, as in *habla-*), but one cannot evaluate this possibility without isolating tense as a variable. For this reason, we focus on three tenses in our case studies: present, preterite, and imperfect.

We adopt Oltra-Massuet and Arregi's (2005) morphological template of the Spanish verb, where tense and agreement are fused in the present and preterite tenses ($\sqrt{-}\text{-Th-T/Agr}$) but not in the imperfect ($\sqrt{-}\text{-Th-T-Agr}$). In the present tense, there are contrasts between null

and overt morphology for both person and number that align with the notion of underspecified versus specified features. Specifically, the null morpheme that encodes third-person singular contrasts with the *-o* that encodes first-person singular and the *-n* that encodes third-person plural.³

(16) Present (Person)

- a. habl- \emptyset -o [1]
- b. habl-a- \emptyset []

(17) Present (Number)

- a. habl-a-n [PL]
- b. habl-a- \emptyset []

In the preterite tense, the contrast is inverted for person: a null morpheme spells out first person and *-ó*, third person.⁴ For number, the contrast is between two overt morphemes: *-ron* for plural and *-ó* for singular.

(18) Preterite (Person)

- a. habl-é- \emptyset [1]
- b. habl- \emptyset -ó []

(19) Preterite (Number)

- a. habl-a-ron [PL]
- b. habl- \emptyset -ó []

Lastly, in the imperfect tense, the only contrast between null and overt morphology relates to number.

(20) Imperfect (Person)

- a. habl-a-b-a- \emptyset [1]
- b. habl-a-b-a- \emptyset []

(21) Imperfect (Number)

- a. habl-a-b-a-n [PL]
- b. habl-a-b-a- \emptyset []

In short, third-person singular as a pairing of default feature values is [] across these tenses, whereas the morpheme that encodes these values is null in the present and imperfect tenses but overt in the preterite. To our knowledge, no study has considered this distinction when exploring the acquisition of agreement morphology in Spanish or in any other language.

If L2 errors stem from the substitution of underspecified feature values for specified ones—and not from the exponents that realize these values—the frequency of errors should be similar in all three tenses. This outcome would strengthen McCarthy’s (2006, 2012) proposal by pinpointing the error source to a moment in the derivation that precedes lexical insertion. If, on the other hand, L2 errors stem from the substitution of morphophonologically simple forms, the frequency of errors should be higher in tenses where default feature values correspond to overt forms, such as the preterite. This outcome would support Bruhn de Garavito’s (2003b) suggestion that learners tend to substitute verbal stems in Spanish. We now present the results of each case study to evaluate these alternatives.

³For reasons of space, we only include the first conjugation (*-ar*), but these patterns are representative of the second (*-er*) and third (*-ir*) conjugations as well.

⁴Even if one does not subscribe to the segmentation in (18a) and instead assumes that *-é* is an overt expression of first-person singular agreement, the contrast between third-person and first-person singular in the preterite tense is still different from the contrast between these forms in the present.

5.1 Person

For person, we compare substitution rates in the present tense, where third-person singular agreement is null, and in the preterite, where third-person singular agreement is overt (-*ó*). We do not include the imperfect in this case study due to the syncretism between first-person and third-person singular forms, where agreement is null.

Among the present tense targets in our data set, there are 227 first-person singular targets and 214 (94%) productions in the target present tense. Of these 214 tokens, 7 (3%) are third-for-first person substitutions, where a null morpheme is used in place of an overt one. Among the preterite tense targets, there are 209 first-person singular targets and 138 (66%) productions in the target preterite tense. Of these 138 tokens, 35 (25%) are third-for-first person substitutions, where an overt morpheme used in place of a null one. Table 6 summarizes these results.

Table 6. Person (1SG target, 3SG produced form)

	1SG targets	Same tense	3SG substitutions	3SG form
Present	227	214 (94%)	7 (3%)	Null
Preterite	209	138 (66%)	35 (25%)	Overt

The data reveal that, for person, null morphology is not substituted at a higher rate than overt morphology. In fact, we observe the opposite tendency. In the preterite tense, where third-person singular agreement is overt, the rate is 25%, while in the present, where third-person singular agreement is null, the rate is 3%. These findings support McCarthy's (2006, 2012) analysis and run counter to Bruhn de Garavito's (2003b).

5.2 Number

For number, we include the imperfect tense since there is a contrast between null third-person singular agreement and overt third-person plural agreement (-*n*), just as there is in the present. We compare substitution rates in these tenses with those in the preterite, where third-person singular agreement is overt (-*ó*).

Among the present tense targets in our data set, there are 67 third-person plural targets and 58 (87%) productions in the target present tense. Of these 58 tokens, 13 (22%) are singular-for-plural substitutions, where a null morpheme is used in place of an overt one. Among the imperfect targets, there are 56 third-person plural targets and 15 (27%) productions in the target imperfect tense. Of these 15 tokens, 3 (20%) are singular-for-plural substitutions, where a null morpheme is also used in place of an overt one. Lastly, among the preterite tense targets, there are 40 third-person plural targets and 27 (68%) productions in the target preterite tense. Of these 27 tokens, 8 (30%) are singular-for-plural substitutions, where an overt morpheme is used in place of another distinct overt morpheme. All of these calculations are summarized in Table 7.

Table 7. Number (3PL target, 3SG produced form)

	3PL targets	Same tense	3SG substitutions	3SG form
Present	67	58 (87%)	13 (22%)	Null
Imperfect	56	15 (27%)	3 (20%)	Null
Preterite	40	27 (68%)	8 (30%)	Overt

As we found for person, null morphology is not substituted at a higher rate than overt morphology for number. For number, like person, the data show the opposite tendency. In the preterite tense, where third-person singular agreement is overt, the rate is 30%, while in the present and imperfect, where third-person singular agreement is null, the rates are 22% and 20%, respectively. These findings also favour McCarthy's (2006, 2012) proposal over Bruhn de Garavito's (2003b).

5.3 Summary

To determine whether L2 errors are more sensitive to morphological defaults (simplified featural structures in the morphological component) or to morphophonological defaults (less specified vocabulary items, including null exponents), we compared the rate of default substitutions in three commonly used tenses. Doing so provided a means of testing the contrasting predictions that arise from the Missing Surface Inflection Hypothesis, as articulated in Bruhn de Garavito's (2003b) study of L2 Spanish, and McCarthy's (2006, 2012) Morphological Underspecification Hypothesis.

In our case studies, we found that substitution rates are not higher when default feature values correspond to less specified, or null, forms. The results indicate that L2 inflectional errors are not driven by properties of morphophonological exponents, challenging Bruhn de Garavito's (2003b) analysis and reinforcing McCarthy's (2006, 2012). In addition, our findings suggest that the source of L2 errors is the mapping of syntactic terminal nodes to morphological structures, which precedes the mapping of morphological features to vocabulary items.

6. Conclusion

Our corpus study finds that L2 errors in inflectional morphology of the Spanish verb are systematic, showing a clear tendency towards the overproduction of forms that realize the default feature values for tense (present), person (third person), and number (singular). Interestingly, the rate of substitution and the type of overproduced default value varies according to whether the target contains a marked or default value. In light of these results, generalizations regarding L2 inflectional errors can and should be made at the morphological level, namely in terms of feature structures.

Bruhn de Garavito's (2003b) verbal stem hypothesis predicts that errors consist of shorter forms across the board, irrespective of whether the omitted morphemes are exponents of tense, person, or number. On the other hand, McCarthy (2006, 2012) argues that

substitution errors exhibit a pattern of overproducing underspecified forms. However, three issues remain to be clarified. First, the notion of “default” is ambiguous. One possibility is that interlanguage grammars are prone to underspecification at the level of feature structures, and another is that L2 speakers tend to insert less specified, or null, vocabulary items. Second, related to this ambiguity, there are cases where default feature values are not realized by less specified vocabulary items, and the role that default forms may play in these cases has not been considered. Finally, McCarthy’s dichotomy between underspecification errors and feature clash errors does not capture the possibility of marked-for-unmarked value substitutions, such as first person for third person.

Our findings suggest that L2 inflectional errors are modulated by morphological feature structure, and a significant amount of non-target productions originate in the post-syntactic mapping of fully-specified syntactic structures to reduced morphological structures. In many instances, these errors can be analyzed as the outcome of “overpruning” marked features from syntactic terminals, which leads to the overuse of forms corresponding to default, or underspecified, values. Our analysis shows that the L2 tendency to substitute defaults cannot be attributed to less complex forms, thereby ruling out the possibility that errors stem from vocabulary insertion, or the mapping of morphology to phonology. Further research is needed to determine whether the notions of defaults and markedness affect L2 production at the minimally local morphemic level or whether there are systematic morpheme interactions in terms of hierarchical feature structures.

References

- Bruhn de Garavito, Joyce. 2003a. The (dis)association between morphology and syntax: The case of L2 Spanish. In *Linguistic theory and language development in Hispanic languages*, ed. Silvina Montrul and Francisco Ordóñez, 398–417. Somerville, MA: Cascadilla Press.
- Bruhn de Garavito, Joyce. 2003b. Learners’ competence may be more accurate than we think: Spanish L2 and agreement morphology. In *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002)*, ed. Juana M. Liceras, Helmut Zobl, and Helen Goodluck, 17–23. Somerville, MA: Cascadilla Proceedings Project.
- Cowper, Elizabeth. 2005. The geometry of interpretable features: Infl in English and Spanish. *Language* 81(1): 10–46.
- Cuervo, María Cristina and Gabrielle Klassen. 2014. Featural (under)specification in the L2 acquisition of tense. Paper presented at the 24th annual conference of the European Second Language Association (EUROSLA), York.
- Domínguez, Laura, Nicole Tracy-Ventura, María J. Arche, Rosamond Mitchell, and Florence Myles. 2013. The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism: Language and Cognition* 16(3): 558–577.
- Halle, Morris. 1997. Distributed Morphology: Impoverishment and fission. In *MIT Working Papers in Linguistics 30*, ed. Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, 125–149.
- Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Ken Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: The MIT Press.

- Halle, Morris and Alec Marantz. 1994. Some key features of Distributed Morphology. In *MIT Working Papers in Linguistics 21*, ed. Andrew Carnie and Heidi Harley, 275–288.
- Harley, Heidi and Rolf Noyer. 1999. Distributed Morphology. *Glott International* 4(4): 3–9.
- Harley, Heidi and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3): 482–526.
- Herschensohn, Julia. 2001. Missing inflection in second language French: Accidental infinitives and other verbal deficits. *Second Language Research* 17(3): 273–305.
- Lardiere, Donna. 1998a. Case and tense in the ‘fossilized’ steady state. *Second Language Research* 14(1): 1–26.
- Lardiere, Donna. 1998b. Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research* 14(4): 359–375.
- Lardiere, Donna. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25(2): 173–227.
- Marantz, Alec. 1997. No escape from syntax: Don’t try morphological analysis in the privacy of your own Lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2): 201–225.
- Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1): 69–109.
- McCarthy, Corrine. 2006. Default morphology in second language Spanish: Missing inflection or underspecified inflection? In *New perspectives on Romance linguistics, vol. 1: Morphology, syntax, semantics, and pragmatics*, ed. Chiyo Nishida and Jean-Pierre Y. Montreuil, 201–212. Amsterdam: John Benjamins Publishing Company.
- McCarthy, Corrine. 2012. Modeling morphological variation and development: Person and number in L2 Spanish. *Linguistic Approaches to Bilingualism* 2(1): 25–53.
- Meisel, Jürgen M. 1997. The acquisition of the syntax of negation in French and German: Contrast-ing first and second language development. *Second Language Research* 13(3): 227–263.
- Montrul, Silvina. 2011. Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition* 33(2): 163–192.
- Montrul, Silvina. 2021. Morphology in Spanish heritage language grammars. In *The Routledge handbook of Spanish morphology*, ed. Antonio Fábregas, Víctor Acedo-Matellán, Grant Armstrong, María Cristina Cuervo, and Isabel Pujol Payet, 538–549. New York, NY: Routledge.
- Oltra-Massuet, Isabel and Karlos Arregi. 2005. Stress-by-structure in Spanish. *Linguistic Inquiry* 36(1): 43–84.
- Perpiñán, Silvia. 2021. Morphology and L2 acquisition. In *The Routledge handbook of Spanish morphology*, ed. Antonio Fábregas, Víctor Acedo-Matellán, Grant Armstrong, María Cristina Cuervo, and Isabel Pujol Payet, 526–537. New York, NY: Routledge.
- Prévost, Philippe and Lydia White. 2000. Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16(2): 103–133.
- Spanish Learner Language Oral Corpora. 2020. <http://www.splloc.soton.ac.uk>. Accessed July 2020-May 2021.
- Vainikka, Anne and Martha Young-Scholten. 1994. Direct access to X’-theory: Evidence from Korean and Turkish adults learning German. In *Language acquisition studies in generative grammar*, ed. Teun Hoekstra and Bonnie D. Schwartz, 265–316. Amsterdam: John Benjamins Publishing Company.
- Vainikka, Anne and Martha Young-Scholten. 1996. Gradual development of L2 phrase structure. *Second Language Research* 12(1): 7–39.