Underspecification and Default Morphology in Second Language Spanish

Corrine McCarthy
McGill University

1. Introduction

Several theories have attempted to explain the lack of consistent suppliance of second language (L2) morphology. The discussion has focused on the cause of these errors: do they follow from impaired syntax (Vainikka and Young-Scholten 1994; Hawkins and Chan 1997) or mapping problems between syntax and morphology (Lardiere 1998a,b; Prévost and White 2000)? On both accounts, the notion of default has often been employed in a post-hoc manner in addressing substitutions of one form for another. While it has been observed that learners do employ default forms, the actual morphemes that are employed as defaults are not predicted by any of these theories. These accounts aim to explain the presence of variability, but not the variants themselves.

In the sections that follow, I propose that morphological variability is constrained in predictable ways. In particular, the choice of defaults found in nontargetlike productions is constrained according to principles of underspecification of morphological features. While full-specification is not fatal in itself, current theories fail to provide a principled account of why certain morphemes emerge as defaults. Under the approach advanced here, only underspecified forms may act as defaults; by assumption, independent properties of a grammatical system determine which features are underspecified. Data come from the spontaneous production of non-native speakers of Spanish, and address two domains: (1) person, number, and finiteness in the verbal domain; (2) gender and number in determiners, a nominal domain.

This paper is organized as follows: in Section 2, I discuss recent literature on the use of defaults in L2 morphology. I will show that although the literature evokes the notion of default, a principled account of why particular morphemes emerge as defaults is lacking. In Section 3, I discuss the use of underspecification in theoretical morphology, and present the Morphological Underspecification Hypothesis. This hypothesis predicts that the morphemes employed as defaults are always underspecified (or less specified), thereby predicting which defaults may surface in learner productions. A set of independent criteria is established for determining which features are underspecified in the grammar. In Sections 4 and 5, I present and discuss original data from speakers of L2 Spanish that support the hypothesis presented here. In Section 6, I compare the data gathered here to other studies of L2 morphology, and discuss the implications for a theory of L2 morphological features. I argue that the results presented here support a theory that assumes underspecification.

2. Background: Variability and Defaults in L2 Morphology

While it is clear that the production of inflectional morphology is variable, theories differ as to the source of this variability. On one type of account, it has been proposed that missing morphology derives from impaired syntax. This impairment can include a lack of functional projections, features and/or feature strength (Vainikka and Young-Scholten 1994; Hawkins and Chan 1997, among others). Under this approach, morphology drives syntactic verb raising: it should never be the case, therefore, that a verb raises to a higher functional projection in the absence of inflection. This approach to (native) morphosyntax has been elaborated under the Rich Agreement Hypothesis (Rohrbacher 1999).

Evidence against the impairment-based approach to non-targetlike L2 morphology comes from facts about word order. Prévost and White (2000) have shown that impaired morphological production does not entail an absence of verb raising or the functional projections (IP/AgrP) which are associated with agreement. Prévost and White (2000) found that, in L2 French and German, learners sometimes raise uninflected verbs over adverbs and negation. Extending the argument from word order to case, Lardiere (1998b) showed the absence of overt tense morphology does not entail the absence of TP, if this is assumed to be the locus of nominative case assignment; Lardiere's L2 English speaker exhibited perfect nominative case assignment, but impaired tense morphology. Together, these results suggest that the absence of agreement and tense morphology in L2 does not imply syntactic impairment. Furthermore, typological evidence contradicts the theoretical claim that overt morphology drives verb raising: languages are found in

which verb raising is allowed, but morphology is impoverished (e.g. Afrikaans), thereby implying that the relationship between overt morphology and verb raising is not bidirectional (Bobaljik 2001). Together, the facts about L2 morphology and native-language morphology indicate that a lack of morphology does not imply a lack of syntactic projections that can serve as the landing site for verb raising.

Prévost and White's and Lardiere's findings argue in favor of another model of L2 grammar, in which morphological variability derives from a mapping problem between syntax and morphology, with functional projections and features intact: this has been proposed under the Missing Surface Inflection Hypothesis (MSIH; Prévost and White 2000, see also Haznedar and Schwartz 1997, Lardiere 2000). Furthermore, the MSIH presents an underspecification-based approach, albeit limited to the underspecification of finiteness. Spontaneous production data from L2 French and German speakers at low proficiency levels showed that non-finite forms (that is, those lacking overt tense/agreement morphology) sometimes occurred in finite (raised) positions. The reverse did not occur: finite verbs were generally limited to finite positions, and did not occur in nonfinite contexts. Nonfinite forms, they suggest, act as defaults because they are underspecified for finiteness, bearing the specification [α finite].

Prévost and White further assume that features and feature-checking mechanisms are unimpaired, and therefore predict that if learners produce agreement, it will always be accurate. This was borne out: the overall error rate for regular and irregular verbs was around 5 percent. However, the low error rate for regular verbs is difficult to interpret, as French is homophonous across all persons in singular forms of present tense. As it turns out, this homophony is important: the most common type of person error found in the data collected here is the substitution of 3rd person for 1st. In Spanish, these forms are, for the most part, distinct; in French, the language Prévost and White investigated, they are usually homophonous. French irregular verbs *etre*, *avoir* and *aller* did surface with a few agreement errors, mainly involving the overgeneralization of 3rd person singular forms to other persons:

1. j'a   fait   la bagarre avec lui-meme
   I have-3s made   the fight   with him

In sum, "default", under the MSIH, is equivalent to an absence of inflection. The MSIH makes no predictions regarding variability among finite forms, only to say that inflection, when supplied, is accurate[1].

In nominal domains, several studies have sought to answer the question of whether functional features like gender that are not instantiated in the L1 are acquirable in the L2. Although these studies do not directly seek to answer the question of which gender emerges as a default, they have found that learners do tend to employ one gender as default, and it is typically masculine (White et al 2004, Bruhn de Garavito and White 2002, Franceschina 2001, but see Hawkins 1998 for counterexamples). White et al (2004) examined elicited production data in L2 Spanish (L1 English, L1 French) for evidence of gender morphology. Overall, correct gender agreement on determiners was produced at a rate of 83 to 99 percent, depending on proficiency level[2]. Furthermore, the L1 English group showed no significant difference in accuracy from an L1 French group, despite the presence of gender as a functional feature in French. White et al conclude that functional features such as gender that are not instantiated in the L1 are acquirable in the L2. Upon closer investigation of those errors that did occur, learners were significantly less accurate overall at producing feminine agreement than masculine agreement. The lower rate of accuracy on feminine determiners was an effect of the substitution of masculine for feminine determiners; participants rarely substituted feminine determiners for masculine ones. (Number accuracy was generally higher, although the authors did not report the defaults learners employed, whether singular or plural.) White et al suggest that masculine is a default following theoretical literature on Spanish (Harris 1991; see also Bruhn de Garavito and White 2002).

In a case study of one L2 end-state Spanish speaker, Franceschina (2001) found that, with a few

---

1 Prévost and White briefly discuss the notion that 3rd person may be underspecified, following a suggestion made by Ferdinand (1996) for L1 acquisition. However, the MSIH does not predict this result. Underspecification of 3rd person would be at odds with their suggestion that inflection, when supplied, is accurate. If Prévost and White predict that inflection in person agreement is accurate (contra the impaired-syntax view), then they cannot also predict an asymmetry across persons.

2 White et al separate their results into Det N and Det N Adj contexts; here, I report only the results from Det N contexts. A significant difference between contexts was found in that speakers were worse overall when adjectives were produced. In the present study, I do not make a distinction between contexts.

exceptions, masculine was the choice of default: this held true for determiners, adjectives, pronouns, and demonstratives. Rather than attributing this tendency to underspecification (and contra White et al 2004, Bruhn de Garavito and White 2002, and the proposal I will advance here), Franceschina attributes it to incorrect or absent specification of gender in the syntax. Franceschina further assumes full specification with respect to gender, meaning that both [masculine] and [feminine] features are equally available in the lexicon. This assumption means that there is no principled reason for masculine to act as default.

To summarize, current full-specification L2 theories offer no account for why certain defaults emerge. In principle, these theories could just as easily derive 2nd person, plural, or feminine defaults— *none* of which surface as defaults— as 3rd person, singular, and masculine ones, all of which *do* surface. In Section 3, I elaborate the theoretical assumptions that derive default morphology, and show how the morphemes that surface as defaults are predictable. While this paper is not the first to suggest that L2 defaults are instances of underspecified morphology, it is perhaps the first attempt to predict what types of morphological errors learners make, and to unite facts about default morphology across nominal and verbal domains under a single unifying principle. This means that much of the previous literature has missed an important generalization about what can and cannot be a default, and what features defaults can realize.

There are two points at which the methodology of the present study diverges from the previous studies of L2 morphology discussed here. First, the spontaneous production studies conducted to date have been case studies of one particular L2 speaker (Lardiere 1998a, 1998b, Franceschina 2001) or have had a very small number of participants (Prévost and White 2000). The present study examines production data from eleven L2 speakers, and can be interpreted as more representative of L2 Spanish speakers in general. Second, the level of proficiency of participants in the present study ranges from intermediate to advanced, whereas the level of proficiency in previous studies has been either end-state (Lardiere)/near-native (Franceschina) on one hand, or quite low on the other (Prévost and White). Little is known about the productions of L2 speakers at intermediate levels. This study also avoids a complication of some earlier studies (Lardiere, Prévost and White) that looked for evidence of knowledge of finiteness and agreement in languages that either lack unique nonfinite bound morphology (English) or show extensive homophony (English, French). By turning to Spanish, which has morphologically-identifiable nonfinite forms and relatively little homophony, a more detailed analysis of finiteness and default morphology is possible.

In the following section, I discuss the principles of underspecification that underlie the theory of interlanguage morphology presented here, and state the predictions of the Morphological Underspecification Hypothesis. These principles allow a precise definition of the notion of default, something that is lacking in current L2 literature.

3. Background: Underspecification
3.1 Underspecification in Theoretical Morphology

Theories of underspecification emphasize economy of representation by excluding redundant information. Arguments in favor of underspecification cite its predictive capacity regarding the shape of inventories, the restriction of combinatorial possibilities available to the grammar, and the neutralization of contrast under certain conditions. I further assume that underspecified forms correspond to unmarked forms, following assumptions made in theoretical literature on morphological features (see Noyer 1992, Harley and Ritter 2002, Carstairs-McCarthy 1998, and many others). In this section, I present independent criteria for establishing markedness values for the variables under consideration, and make reference to the theoretical literature that adopts these values as unmarked.

For the variable of person, there is some evidence to suggest that 3rd person is universally unmarked. Greenberg (1966) notes that unmarked values tolerate more distinctions than marked ones. Typologically, 3rd person is more likely than 1st or 2nd person to show gender/number distinctions, suggesting that 3rd person is unmarked. Harley and Ritter (2002) adopt a feature-geometric approach to person-number features, and use typological evidence to motivate their geometry. Markedness values are reflected in their geometry, in that unmarked values have fewer nodes. Under their approach, 3rd person and singular are underspecified: 3rd person is realized via the absence of a participant node; singular is realized via the presence of a bare individuation node and absence of a group node. For verbal morphology relating to finiteness, it is further assumed that nonfinite forms are underspecified for finiteness, as they lack specification for tense (Jakobson 1984).

Turning to gender, there is some evidence to suggest that masculine is unmarked. One approach to

determining which form is unmarked is the criterion of neutralization; this was originally developed as a criterion under structuralist phonology. Neutralization occurs when a marked term is excluded from some context in which an unmarked term can occur (Battistella 1990). In Spanish, there are many pairs of words of which one member is specified for feminine (*hermana* 'sister'), while the other member is apparently specified for masculine (*hermano)*. If these are pluralized, the category *hermanos* can occur in the context of describing a group of male and female siblings; the feminine *hermanas* cannot occur in such a context. The broader distribution of *hermanos* over *hermanas* suggests that *hermano* is unmarked. On the assumption that unmarked values should be underspecified, I assume that [masculine] is underspecified.

Harris (1991) has argued, on the basis of facts about the Spanish lexicon, that masculine is underspecified in Spanish. One example that supports his claim is the use of the preposition *para* 'for':

2. Tienes   demasiados   "paras" en este parrafo   (Harris 1991:43)
      Have-2sg too-many-masc paras   in this paragraph.

Since *para* is a preposition, it has no gender and cannot transfer gender to the quantifier *demasiados*. There is no alternative source of masculine gender in this sentence, so this presents a clear argument in favor of masculine gender as the default, at least in the case of Spanish.

In this section, a theoretical basis has been established for predicting which forms will act as defaults. Following generalizations set out in the theoretical literature, 3rd person, masculine, and singular are unmarked, and by assumption, unavailable as features in the L2 lexicon. For the variables considered here, underspecification means that there is an asymmetry built into the feature specifications of morphemes, whereby some morphemes are associated with more features than others. This asymmetry distinguishes the current proposal from full-specification theories by making it possible to predict which morphemes will emerge as defaults. The precise mechanisms that yield these defaults are elaborated in Section 3.2.

3.2 Deriving Defaults

I assume Distributed Morphology, a theory that adopts underspecification and late-insertion of vocabulary items. I further assume privative feature values: features are present or absent, with absent features corresponding to unmarked values. For example, for the variable of gender, the absence of [feminine] implies masculine gender; masculine is represented by the absence of a gender feature. Vocabulary insertion is a competition in which the most highly specified vocabulary item, barring feature clash, is inserted into the fully-specified syntax. Where no features match between the terminal node and feature bundle, an elsewhere form is inserted. Competition for vocabulary insertion proceeds from the most highly specified entry to least specified entry (the elsewhere form). Where there is an equal number of features specified for two or more forms, the order must be stipulated (Halle and Marantz 1993).

Vocabulary insertion involves the competition for insertion into an abstract AGR morpheme in the syntax. Vocabulary items that compete for insertion are listed in (3), limiting discussion to present tense, *-ar* class[3].

3.   AGR
       [1][plural]   ↔   -amos
       [plural]   ↔   -an
       [1]   ↔   -o
       [2]   ↔   -as
       elsewhere ↔   -a

The vocabulary insertion rules in (3) yield the paradigm for Spanish person-number agreement in the present tense, given in (4). This paradigm is representative of most dialects of Latin American Spanish.

---

3 For ease of presentation, this paradigm represents only the *-ar* class in present tense. An additional mechanism is necessary to distinguish conjugation class, but will not be discussed here. I limit discussion to present tense in order to avoid making claims on the specification of tense features at this point.

4. Infinitive: hablar 'speak/talk'
      1sg hablo   1pl hablamos
      2sg hablas   2pl hablan
      3sg habla   3pl hablan

In order to illustrate how vocabulary insertion operates, consider the syntactic context corresponding to 1st person singular *yo*: an AGR node specified for [1][singular]. In (5), a fully-specified syntactic terminal node (5a) interfaces with the vocabulary items (5b). In accordance with the notion that 3rd person and singular are underspecified, [3] and [singular] are not available as features in (3/5b): only [1], [2], and [plural] are available. Competition for vocabulary insertion will bar the most highly-specified vocabulary item *-amos*; this item realizes [1][plural], and since [singular] clashes with [plural], it is not inserted. For the same reason, *-an* is exluded. *-o* matches for the feature [1], and yields no feature clash. This vocabulary item therefore wins the competition for insertion.

5a. syntactic terminal node   b.   vocabulary items
         AGR   [1][plural] ↔ -amos
         [1][singular]   [plural]   ↔ -an
                                                [1]   ↔ -o
                                                [2]   ↔ -as
                                                elsewhere ↔ -a

In a syntactic context of 3rd person singular, the only option is the elsewhere morpheme *-a*; any other vocabulary item will yield a feature clash of person and/or number.

Turning to determiner morphology, Spanish nouns trigger agreement in gender and number. The same process of vocabulary insertion applies to gender. I assume an abstract morpheme D, the head of DP, and the feature specification of vocabulary items listed in (6)[4,5] (limiting our discussion to definite determiners):

6.   vocabulary items
      [feminine][plural] ↔ las
      [plural]   ↔ los
      [feminine]   ↔ la
      elsewhere   ↔ el

As stated above, the unmarked features of [masculine] and [singular] are underspecified. Only the features [feminine] and [plural] are available.

In principle, two types of errors may occur in the competition for lexical insertion: feature clash and underspecification. For person agreement, [3] is assumed to be underspecified. Suppose the syntax supplies the [1] feature, but instead *-as* is produced, a realization of the [2] feature:

7. yo hablas
    I   speak-2sg

This is an error of feature clash. Suppose, on the other hand, that the syntax supplies the [1] feature, but instead *-a* is produced, a realization of the elsewhere condition:

8. yo habla
    I   speak-3sg

---

4 The precise location of features is not immediately relevant. White et al (2004) assume these features to be present in the head of NumP, a projection between DP and NP, in L2 Spanish.
5 I also examine indefinite determiners. The same feature specifications apply to the indefinites. I do not include both definites and indefinites in one hierarchy because I do not wish to make a claim about the competition between definite and indefinite determiners at this point.

This is an error of underspecification: where the more highly specified form *-o* should have won the competition for vocabulary insertion, the underspecified form was chosen instead. This does not result in feature clash, since the elsewhere morpheme represents an absence of features. Extending this pattern from person to number, the substitution of a singular form for a plural one constitutes an underspecification error[6].

For the category D, the same logic applies: if the syntax supplies the feature [masculine], the insertion of a feminine form results in feature clash:

9.　la　　　　libro
　　DET-fem book (masc, sg)

The insertion of a masculine determiner in a feminine context results in an error of underspecification, as the elsewhere morpheme *el* represents an absence of gender features:

10.　el　　　　noche
　　DET-masc   night (fem, sg)

Similarly, the occurrence of a nonfinite form in a finite context would constitute an underspecification error. The occurrence of a finite form in a nonfinite context would constitute a feature clash.

It is hypothesized that learners will avoid feature clash in their interlanguage grammar. When errors occur, they are errors of underspecification. I propose the hypothesis in (11):

11. Morphological Underspecification Hypothesis: L2 errors are ones of underspecification, not of feature clash.

This prediction is tested for the following variables in verbal and nominal domains:

　　a.　Verbal domain: Person, number, and finiteness
　　b.　Nominal domain: Gender and number in determiners

4. Methodology
4.1 Data collection

The data come from spontaneous production of speakers of Spanish as a second language. There were eleven participants included in the data set, all of whom began learning Spanish after age 12. All participants were asked to rate their level of proficiency in spoken Spanish. Responses ranged from intermediate to advanced, except for one who reported being near-native. Participants also completed a proficiency test that consisted of a cloze test and a multiple choice vocabulary/grammar test; all participants scored either intermediate or advanced. The majority (nine of eleven) of participants had received both naturalistic and classroom exposure to Spanish: they reported having lived in a Spanish-speaking environment for four weeks or more. Two participants reported never having used Spanish outside the classroom. All had received at least one semester of formal instruction. There were 15 participants initially, but 4 were excluded. One was excluded since his data yielded no errors in any of the variables of interest. The other three were excluded due to French exposure during the critical period; since the research conducted in the primarily French-speaking city of Montreal, it was difficult to find true cases of Spanish as a second (not third) language. It was decided that pre-critical period exposure to a Romance language might introduce unwanted variables, since French has similar properties to the L2 target language, Spanish.

Interviews were conducted by a native speaker of Spanish. Participants were told that they should consider the interview a "casual conversation", and were encouraged to ask questions of the interviewer if

---

6　While I focus on the use of 3rd singular as a default, it should be noted that these are not the only kind of underspecification errors that are possible under the hypothesis I advance here. For example, 1st singular may substitute for 1st plural without any feature clash; 3rd plural may substitute for 1st plural without any feature clash. As it turns out, these type of errors do not occur. This issue is further addressed in Section 6.

they wanted. The interviewer had a predetermined set of topics and questions, but she was encouraged to allow the participants to talk about any topic that interested them in order to elicit the most naturalistic speech possible. Topics of discussion frequently included travel abroad, academics, family, and daily life in Montreal. Interview lengths ranged from 15 to 35 minutes.

Speech was transcribed by a near-native speaker of Spanish. Following the methodology of Lardiere (1998a,b), utterances that were followed by self-correction were excluded; the final, corrected forms were included. Self-repetitions and repetitions of the interviewer were excluded. Errors were coded for type according to the criteria described below.

a) AGR morphology

All errors in person and number agreement were coded for whether agreement constituted an error of feature clash or of underspecification. Finite forms in which the underspecification of [3] and [singular] yielded no feature clash were classified as underspecification errors[7]. Nonfinite forms (infinitives) that occurred in finite contexts were classified as underspecification errors. As Spanish permits null subjects, only those verbs whose (null or overt) subject was judged to be unambiguous were included in the analysis. This was done in order to eliminate any chance of reporting an error when none was produced.

b) Determiners

All errors in gender and number were coded for whether the produced form constituted an error of feature clash or underspecification based on the underspecification of [masculine] and [singular]. Plural determiners occasionally contained a reduced vowel, which made it difficult to determine whether the target form was *los* or *las*, *unos* or *unas*. These tokens were excluded.

5. Results.
5.1 AGR morphology

A total of 62 errors in verb agreement were found. Error types are presented in Table A. Excluding data from one participant, 88 percent of errors are underspecification errors. One participant repeatedly employed a first-person past preterite copula/auxiliary, *estuve* 'was' , in 3rd person contexts (see Section 6), constituting an error of feature clash. She does not, however, use first person as a default in other tenses: in simple present, 3rd person surfaces twice in a non-targetlike context.

Table A. Person agreement: Number of tokens by participant and error type

	Participants 1-10	Participant 11
Underspecification Errors	14 (88%)	2 (10%)
Feature Clash Errors	2 (12%)	19 (90%)
Total Errors	16	21

(12,13,14) below are examples of errors of underspecification in person agreement. In (12), the participant asks a question that lacks 2nd person agreement. There was a slight pause between repetitions of the verb, as she was waiting for the interviewer to answer her. The interviewer didn't understand her the first time, since the verb's agreement indicated 3rd person. (13,14) were uttered in contexts in which the participant was asked for information about his or herself, making the intended referent 1st person. (14) includes a self-"correction" that actually replaces the target form with an underspecified one:

12. y **manejará**? (pause) **manejará** tú?   (Samantha, intermediate)
    and   drive-fut-3sg   drive-fut-3sg you
    'and will you drive?'

---

7 Spanish shows homophony between 2nd person singular formal *Usted* and 3rd person singular in all tenses. No participant used *Usted* during the interviews. There were 3 instances in which a verb corresponding to 2nd *tú* was missing the expected suffix. In one of these, *tú* was overtly produced. For the other two, there was no overt pronoun. However, since the participants repeatedly used *tú* throughout the data set, it was assumed that these forms were missing the 2nd person suffix, and that participants were not switching to formal *Usted*. All three tokens were included as underspecification errors.

13. **nació** en Boston   (Beth, intermediate)
    be-born-past-3sg in Boston
    'I was born in Boston'
14. ahora vivo, **vive** en Westmount   (David, intermediate)
    now live-1sg, live-3sg in Westmount
    'now I live in Westmount'

Of the two errors of feature clash, one involved the verb *gustar* 'like/please'; this error is given in (15). This verb is one of a class of psych verbs that requires an experiencer and theme. In Spanish, the experiencer is case-marked dative and does not trigger agreement, suggesting it is an indirect object; however, it behaves like a subject in its ability to control PRO in adjunct clauses (see Montrul 1998 and references therein). Below, the 1st person singular experiencer occurs phrase-initially and is case-marked dative by the presence of 'personal a', a preposition-like morpheme used before human direct and indirect objects. Agreement corresponds to the theme *los gatos*:

15. A mí   me   gustan   los gatos
    a 1sg-obj 1sg-obj like-3pl   the cats
    'I like cats'

The error of feature clash involved agreement with the experiencer rather than the theme, which is a nonfinite clause, and should trigger 3rd singular agreement (16):

16. Los chicos les   **gustan**   pegarse   bien   (Annie, advanced)
    the boys 3pl-obj like-3pl stick-refl   well
    'The boys like to stick themselves to you'

This error is alternatively analyzed a case of agreement with the wrong argument, rather than incorrect agreement. Under this analysis, the speaker has incorrectly produced agreement that corresponds to the experiencer, rather than the theme. The lack of "personal *a*" marker before the dative experiencer *los chicos* suggests that Annie might be treating the experiencer as a "normal" nominative subject that triggers verb agreement. If this analysis is correct, this error is not an error of feature clash.

For number agreement, results are presented in Table B[8]. Underspecification errors account for 92 percent of all number errors produced.

Table B. Number agreement on verbs: Error type

	N
Underspecification Errors	23 (92%)
Feature Clash Errors	2 (8%)
Total Errors	25

Examples of underspecification errors are given in (17,18). (18) involves the use of *gustar*, but in these cases, agreement is not in accordance with the 1st singular experiencer *me*. This is in contrast to the example in (16), an error of feature clash that I proposed to exclude due to agreement with the wrong argument. In the examples below, the target form is 3rd plural.

17. Los   italianos   **puede** entender   un poco   (Linda, advanced)
    the   Italians   can-3sg understand a   little
    'The Italians can understand a little (Spanish)'
18. hay   varias   regiones en el norte que me   **gustó**   (Steve, advanced)
    there-are various   regions   in the north   that 1sg-obj like-past-3sg
    'There are various regions in the north that I liked'

---

8 There were a few tokens of plural agreement with a singular noun *gente*. These were not coded as errors, since *gente* sometimes occurs with plural agreement in the speech of native Spanish speakers.

The generalization with respect to *gustar* appears to be that speakers do use underspecified agreement, but that the presence of a dative experiencer sometimes triggers nontargetlike agreement. Errors with this particular verb continue to surface even in the speech of some of the advanced L2 speakers (Steve and Annie).

For the variable of finiteness, only four instances of nonfinite forms in finite contexts were found in the data set. An example is given in (19). There were no instances of finite verbs in nonfinite contexts.

19. yo nunca **hacer**   los platos   (Beth, intermediate)
    I   never do-inf   the dishes
    'I never do the dishes'

5.2 Determiners

A total of 97 agreement errors were found for determiners. The data for determiners show more errors than for person-number agreement overall. 85 percent of errors were underspecification errors. Of total errors, only three showed a number error: all of these involved the occurrence of a singular determiner in a plural context: for that reason, the two variables of gender and number are collapsed.

Table C. Gender/number agreement in determiners: Error type

	N
Underspecification Errors	82 (85%)
Feature Clash Errors	15 (15%)
Total Errors	97

Underspecification errors included errors in indefinite (20) and definite (21) determiners.

20. No quiero   contestar, pedirlo por **un**   otra   letra   (Samantha, intermediate)
    neg want-1sg answer,   ask it   for det-indef-masc other-fem letter-fem
    'I don't want to ask him for another letter'
21. Hay   **el**   sangre   *acadian* en mi cuerpo   (David, intermediate)
    there-is det-def-masc   blood-fem Acadian in my body
    'There is Acadian blood in my body'

A particularly interesting example is (22), in which the interviewer introduces an unfamiliar word. Samantha does not assign the word a feminine determiner, despite the fact that it ends in -*a*:

22. Samantha: No   tenemos el,   los,   cómo se dice *candle*?
                neg have-1pl det-masc det-masc-pl how   say candle
                'We don't have the, the, how do you say candle?'
    Interviewer: Velas.
    Samantha: **Los**   velas.
              det-masc   candles-fem

Errors with nouns that follow the "canonical" gender patterns in Spanish (feminine nouns end in -*a*, masculine ones in -*o*) are surprisingly frequent. A clear pattern emerges with respect to these canonical nouns: there are many (about 30) instances of canonical feminine nouns occurring with a masculine determiner (e.g. *un mezcla, un palabra, el revista, los islas*), but there are only six instances of a canonical masculine noun occurring with a feminine determiner (*una cuarto, una método, la derecho, las edificios, una camino, una mercado*). This parallels the (native) Spanish lexicon: Harris (1991) notes that there is only one non-exotic noun that ends in -*o* but is feminine (*la mano* 'hand'; another example is *modelo* 'model'); however, there are nearly 600 instances of nouns that end in -*a* but are masculine.

A common error of feature clash involves these exceptional masculine nouns that end in -*a (programa, sistema)*. These account for six of the fifteen errors of feature clash, and can be attributed to an overgeneralization of the "-*a* nouns are feminine" strategy.

6. Discussion

The results presented here support the hypothesis that L2 learners produce errors of underspecification, while avoiding feature clash. This was shown for person, number and finiteness in verbal morphology, and gender and number agreement in determiners. The errors that learners produced in these domains are consistent with the hypothesis proposed here, barring one exceptional participant.

Focusing on verbal morphology, if we compare the predictions of the MUH with those of the MSIH, we see that the MSIH can only account for four of 62 (6 percent) errors of verbal morphology; these are the ones that involved the substitution of a nonfinite form for a finite one. The MUH, on the other hand, can account for 39 of 62 (63 percent) errors, or 37 of 41 (90 percent) if Sheila's repeated use of preterite irregular 1st person verbs are excluded. More errors involve substitution among finite forms than between finite and nonfinite forms, suggesting that the MSIH, though it correctly predicts nonfinite defaults, does not go far enough in terms of predicting default morphology overall.

The results of other studies of L2 morphology are largely consistent with the predictions made here, as discussed in Section 2 (especially with regard to gender; see White et al 2004). The predictions for verbal morphology might, at first, appear to be at odds with the results reported for L2 English in Lardiere (1998b). There, she found that her subject Patty consistently failed to produce 3rd person singular *-s*. If 3rd person is underspecified, as I propose, it might be expected that this form acts as the default: *-s* would appear in 1st and 2nd person environments. Instead, what appears to be the 1st/2nd person form is found in 3rd person contexts, as in (23):

23. My mom also speak Cantonese   (Lardiere 1998b:368)

This finding is not, however, necessarily problematic for the MUH. Although the syntactic context of the verb is finite, the (potentially finite) form *speak* is homophonous with the nonfinite form *speak*. The example in (23) could therefore be analyzed as an instance of an underspecified nonfinite verb, which lacks person-number features entirely. Under this analysis, it would not be incorrectly specified for person-number features, and therefore would not be an instance of feature clash; it is therefore possible to account for (23) under the MUH.

If we assume that Patty's productions of verbs like *speak* in (23) are morphologically nonfinite, this raises the question of why Patty produces so many nonfinite forms in finite contexts, and why these L2 Spanish participants produce so few. It could be that L2 Spanish learners recognize that Spanish, unlike English, is a language with rich agreement, and therefore learners recognize that verbs generally need to bear overt inflection. Bruhn de Garavito (2003) makes a similar suggestion in noting that L2 Spanish/L1 English speakers, even at low levels, generally produce (correct) verb agreement[9]. English learners, on the other hand, may recognize that English does not have rich agreement and that English verbs generally do not need to bear overt inflection, and have trouble recognizing the contexts in which overt inflection does occur.

One clear counterexample to the predictions made here is Sheila's repeated use of 1st person preterite *estuve* 'was' (infinitive *estar*) in 3rd person contexts (24,25). In both examples below, Sheila uses the wrong copula; *ser* would be used in this context, not *estar*. She has a tendency to use *estuve* anywhere she means *was*.

24. el objetivo   **estuve** ...   (Sheila, intermediate)
      the objective was-1sg
      'the objective was...'
25. no **estuve**   peligroso   (Sheila, intermediate)
      neg was-1sg   dangerous
      'it wasn't dangerous'

Estuve occurs 18 times in 3rd person contexts, and never in 2nd person contexts. There is also one instance of the verb *hacer* used in the same way:

9 Bruhn de Garavito (pc) reports that her learners used 3rd singular defaults, although this is not reported in her 2003 paper.

10

26. hay   una que mi novio   ya   **hice**   (Sheila, intermediate)
there-is one that my boyfriend already did-1sg
'There is one that my boyfriend already did'

Although this is clearly a problem for the MUH, it should be noted that Sheila uses 3rd person as a default in present tense. In addition, these forms are irregular past preterite forms, and Sheila shows no evidence that 1st person generally acts as a default in the past preterite. One explanation for her use of *estuve* and *hice* might be that the target form ends in -*o (estuvo, hizo)*, which is the same ending as 1st singular present. It could be that some kind of blocking effect is applying, in that she knows that 1st person morphology doesn't belong there, and the only other form available in her lexicon is one that ends in -*e*. It might be relevant that Sheila has had the least amount of classroom instruction of all seven participants, which might be a factor in why her errors do not resemble those of the other participants. The MUH makes no predictions about the interaction between style of instruction and type or rate of errors, although there may be some relationship among these variables.

I have proposed here that 3rd person can act as a default in non-3rd contexts. However, the theory allows a range of other substitutions that qualify as underspecification errors, but that do not occur in the data (see footnote 6). In general, learners do not make errors in verbal number inflection, unless it involves 3rd person singular forms in a 3rd person plural context, as in (17,18). The pattern that emerges for inflected verbal defaults is that learner errors tend to involve the "elsewhere" morpheme, i.e., the totally underspecified finite form, corresponding to 3rd singular. Furthermore, since [singular] is underspecified, the theory predicts learners might substitute 1st singular for 1st plural. This does not occur in the data. What can be said is that underspecification errors tend to be truly underspecified, at least for verbal morphology. Although this does not directly follow from the MUH, it may need to be added as an additional stipulation.

The predictions of this underspecification-based theory can be extended to other domains where unmarked forms would be expected to emerge as defaults. First, consider tense. If it can be independently established that [past] is marked relative to [present], the MUH would predict that present-tense morphology must be underspecified. This leads to the prediction that present tense forms may act as defaults in past contexts, but not the reverse. This prediction can be tested against the MSIH, which predicts that non-finite forms may replace finite ones, but makes no predictions regarding variability among finite forms. The same can be said for mood, where the underspecification of indicative mood should lead to its use as a default in subjunctive contexts.

Finally, the results of this study help to elucidate our knowledge of the structure of L2 morphological features. Previous studies, including Hawkins and Chan (1997) and Franceschina (2001) have assumed either "naive" full specification, or have suggested underspecification as a post-hoc explanation for observations of patterns in default morphology (White et al 2004 for gender, Prévost and White 2000 for 3rd person). As was shown in previous sections, these theories cannot predict the form of defaults that emerge. I have suggested here that defaults are underspecified forms, and that the inventory of morphological features in L2 may derive from underspecification based on markedness. In this way, this study adds to the literature on the role markedness plays in second language acquisition.

7. Conclusion

The Morphological Underspecification Hypothesis correctly predicts the absence of certain kinds of errors in L2 production. It predicts that learners will not produce errors that result in feature clash, but that errors of underspecification may occur. Underspecified features are, by assumption, those features that are unmarked. These predictions were supported: errors produced were primarily underspecification errors. This pattern was found for person, number, and finiteness in verbal morphology and for gender and number morphology in determiners. These patterns are not predicted under full-specification theories. The MUH is able to predict the shape of learner defaults under one broad principle: in interlanguage grammar, learners avoid feature clash, in much the same way they do in their native-language grammar.

References
Battistella, Edwin L. 1990. *Markedness: The evaluative superstructure of language*. Albany, NY: SUNY
    Press.

Bruhn de Garavito, Joyce & Lydia White. 2002. L2 acquisition of Spanish DPs: The status of grammatical features. In A. T. Perez-Leroux & J. M. Liceras (Eds.), *The Acquisition of Spanish Morphosyntax: The L1/L2 Connection,* 151-176. Dordrecht: Kluwer.

Bruhn de Garavito, Joyce. 2003. The (Dis)association between morphology and syntax: The case of L2 Spanish. In S. Montrul & F. Ordonez (Eds.), *Linguistic Theory and Language Development in Hispanic Languages.* Sommerville, MA; Cascadilla Press, 399-417.

Carstairs-McCarthy, Andrew. 1998. How lexical semantics constrains inflectional allomorphy. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1997,* 1-24. Dordrecht: Kluwer.

Ferdinand, Astrid. 1996. *The Development of Functional Categories: The Acquisition of the Subject in French*. The Hague: Holland Academic Graphics.

Franceschina, Florencia. 2001. Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research* 17:3, 213-247.

Greenberg, Joseph H. 1966. *Language universals with special reference to feature hierarchies.* The Hague: Mouton.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In K. Hale & S.J. Keyser, eds., *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger,* 111-176. Cambridge, MA: MIT Press.

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

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

Hawkins, Roger & Cecilia Yuet-hung Chan. 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. 1998. The inaccessibility of formal features of functional categories in second language acquisition. Paper presented at the *Pacific Second Language Research Forum*, Tokyo, March 1998.

Haznedar, Belma & Bonnie Schwartz. 1997. Are there optional infinitives in child L2 acquisition? In Hughes, E., Hughes, M. and Greenhill, A. (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press, 293-306.

Jakobson, Roman. 1984.The structure of the Russian verb. In L. R. Waugh & M. Halle (Eds.), *Russian and Slavic Grammar Studies 1931-1981.* Berlin: Mouton.

Lardiere, Donna. 1998a. Dissociating syntax from morphology in a divergent L2 end-state grammar. *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. 2000. Mapping features to forms in second language acquisition. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory.* Oxford: Blackwell, 102-129.

Montrul, Silvina. 1998. The L2 acquisition of dative experiencer subjects. *Second Language Research* 14:1, 27-61.

Noyer, Rolf. 1992. Features, projections and affixes in autonomous morphological structure. PhD Thesis, MIT.

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

Rohrbacher, Bernard. 1999. *Morphology-Driven Syntax: A Theory of V to I Raising and pro-drop*. Amsterdam: John Benjamins.

Vainikka, Anne & Martha Young-Scholten. 1994. Direct access to X'-theory: Evidence from Korean and Turkish adults learning German. In T. Hoekstra & B. D. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar*, 265-316. Amsterdam: *John* Benjamins.

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