



HAL
open science

Morphological variability in the comprehension of agreement: an argument for representation over computation

Corrine Mccarthy

► **To cite this version:**

Corrine Mccarthy. Morphological variability in the comprehension of agreement: an argument for representation over computation. *Second Language Research*, 2008, 24 (4), pp.459-486. 10.1177/0267658308095737 . hal-00570752

HAL Id: hal-00570752

<https://hal.science/hal-00570752>

Submitted on 1 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Morphological variability in the comprehension of agreement: an argument for representation over computation

Corrine McCarthy *George Mason University*

Received December 2007; revised May 2008; accepted May 2008

Previous accounts of morphological variability disagree over whether its cause is representational or computational in nature. Under a computational account, variability is confined to production; under a representational account, variability extends to comprehension and is qualitatively similar to variability in production. This article presents experimental evidence from the comprehension and production of gender and number agreement in second language (L2) Spanish clitics and adjectives. Intermediate-level participants show variability across comprehension and production; across tasks, masculine defaults are adopted. Advanced-level participants show less variability, although evidence for masculine defaults emerges across tasks. Number agreement proved relatively unproblematic, except in the production of adjective agreement where singular defaults are systematically adopted by intermediate- and advanced-level speakers. The qualitative similarity of variability across comprehension and production supports a representational account; however, previous research disfavours an account based in syntactic deficits. This article argues for a theory of morphological variability that places the representational cause in the morphology, rather than the syntax.

Keywords: Second language morphology, morphological variability, acquisition of inflection, L2 gender, comprehension of inflection, underspecification

I Introduction

Previous research on the second language (L2) acquisition of inflection – particularly the production of markers of tense and agreement – has focused largely on the syntactic correlates of missing inflection. One approach argues that the correlation between morphological variability, as manifested by the inconsistent suppliance of (correct) morphology, is a consequence of impaired syntactic representations (e.g. Clahsen, 1988; Meisel, 1991; Hawkins and Chan, 1997). These syntactic impairments may involve a failure to acquire ‘new’ features, feature strengths or functional projections (such as IP, TP, AgrP) in an L2. Syntactic impairments may be temporary, representative of a particular stage of development (Vainikka and Young-Scholten, 1994; 1996) or permanent (Clahsen, 1988; Meisel, 1991; Hawkins and Chan, 1997). I call this class of accounts a ‘representational account of variability’, as the source of missing or incorrect inflection lies in the faulty representation of syntax.

A complication for representational accounts lies in the observation that morphological deficits do not necessarily entail syntactic ones. Lardiere (1998a; 1998b), for example, shows that morphology may remain ‘missing’ for many years, even though the L2 acquirer (L2er) has long since successfully acquired syntactic properties of the L2 that are associated with particular manifestations of morphology (see examples 1–2 below). Although the informant, Patty, frequently omitted 3rd singular *-s* and past *-ed*, her accuracy in verb movement is native-like, as is her apparent knowledge of abstract syntactic properties (e.g. Case) that depend on the acquisition of higher functional projections. Prévost and White (2000) similarly document cases of L2 French and German in which a speaker fails to inflect a verb correctly, but demonstrates accurate knowledge of verb movement. Together, these pieces of evidence demonstrate that any correlation between syntactic deficits and the correct use of inflectional morphology is far from absolute, contrary to a representational account like the one described above.

The dissociation between syntax and the use of inflection is frequently taken as evidence in favour of the Missing Surface Inflection Hypothesis (henceforth MSIH; Prévost and White, 2000; see also Haznedar and Schwartz, 1997). Under this approach, syntactic representations are present, but variability surfaces due to performance limitations resulting from communication pressure. In other words, missing inflection is seen

as a failure of access, rather than representation: the L2 speaker correctly represents tense and agreement on some level, but fails to produce inflection, or produces it variably, due to heavy processing demands. I will call this explanation a ‘computational account of variability’.

According to the MSIH, production data may seriously underestimate underlying competence. The implicit claim, therefore, is that by accessing data from alternative measures we get a better estimation of the L2er’s competence. In this article, I test Prévost and White’s (2000: 129) assertion that variability arises due to communication pressures on output. If variability is strictly a production-based phenomenon, then eliminating communication pressure should eliminate variability, or at least severely reduce its appearance.¹ The first goal of this article, therefore, is to test whether variability extends to comprehension. If it does, this observation constitutes a challenge for the view that variability is a product of mere performance or production-based limitations, as the MSIH suggests. On the other hand, such an observation would be consistent with a view that deficits are representational in nature. The second goal of this article concerns the nature of comprehension variability: if variability in comprehension is, in fact, attested; is it qualitatively similar to the variability that is attested in production? If it is, this would add support to the claim that variability is, at least in part, a representational phenomenon and not a strictly computational one. The third goal of this article will be to attempt to account for the results within a representational account of morphological variability; as we will see, current theories fail to predict the generalizations that emerge.

This article is organized as follows. In Section II, I begin by establishing some generalizations about the phenomenon of morphological variability. As we will see, a close examination reveals the theoretical significance of morphological features. In Section III, I review the essentials of morphological feature geometries, which I adopt for the analysis of gender and number. In Section IV, I lay out the research questions that are addressed in a set of two experiments on agreement in L2 Spanish clitics and adjectives. Section V describes these two

¹ An implicit assumption made by the MSIH appears to be that communication pressure affects production, but does not affect comprehension. I do not follow in making this assumption. That is not to say, however, that task effects are ruled out; comprehension tasks may in fact contribute less pressure. This point is addressed in Section VIII.

experiments: one establishes the nature of variability in production, and the other tests for the existence of variability in comprehension. Section VI presents the results, and Section VII addresses their theoretical consequences. I argue that a representational account of variability is supported, but not the type of account that places the source of morphological variability in syntactic deficits. Instead, an alternative account that defines variability in terms of morphological representations is proposed in Section VII. By incorporating hierarchical organization and underspecification of morphological features into the L2 grammar, a theory of variability is able to model generalizations that previous theories have been unable to capture.

II Morphological variability: generalizations

One generalization that emerges from the literature on first language (L1) morphological variability is that when morphology is supplied, it is accurate (Poeppl and Wexler, 1993). This generalization has held for much L2 research as well, particularly when it involves L2 English (e.g. Lardiere, 1998a; 1998b; White, 2007; for child L2, see also Paradis, 2005). Examples (1–3) illustrate missing inflection. A Chinese near-native speaker of English omits 3rd singular *-s* in (1) and past-tense *-ed* in (2). Similarly, an L1 Mandarin speaker omits an indefinite determiner in (3).

- | | |
|---------------------------------------|-------------------|
| 1) because he understand better now | (Lardiere, 1998b) |
| 2) he call me last night | (Lardiere, 1998a) |
| 3) and she made phone call to someone | (White, 2007) |

The problem of variability is not always so straightforward. Variability frequently involves an alternation between correct and incorrect morphology (or, errors of commission), rather than present and missing morphology (or, errors of omission).² Substitutions are shown in L2 Spanish and French for person (4), tense (5), gender (6) and finiteness (7). The target form in (4) is *nací* ('I was born'); the error involves substituting 3rd person *-ió* for 1st person *-í*. The target form in (5) is the past preterit *gustó* or imperfect *gustaba* 'liked'; the error involves substituting present tense

² Missing inflection is, unfortunately, an ambiguous term. For some authors, it is used to describe the absence of overt inflection, as in examples (1–3). For others, it means an absence of finiteness, presumably at an abstract level (see Prévost and White, 2000). I take the view that missing inflection under both interpretations is insufficient to describe the range of verbal and nominal morphological variability, as learners use incorrect overt morphology in (4–7) and use finite forms as defaults in (4–5).

gusta for past. The target form in (6) is *roja* ‘red-FEM’; the error involves substituting *-o* for *-a*. In (7), an L2 French speaker substitutes a non-finite form, *arriver* ‘arrive-INF’, for a finite one: *arrive* ‘he arrives’.

- 4) Nació en Boston. (McCarthy, 2007)
 was.born-3SG in Boston
 ‘I was born in Boston.’
- 5) En el pasado, me gusta mucho mi trabajo. (Mezzano, 2003)
 In the past, to-me like-PRES a lot my job
 ‘In the past, I liked my job a lot.’
- 6) la barba rojo (White *et al.*, 2004)
 the-FEM beard red-MASC
 ‘the red beard’
- 7) monsieur il arriver (Prévost and White, 2000)
 mister he arrive-INF

The term ‘default’ has frequently been used to describe the systematic outcome of substitution errors; (4–7), then, can be viewed as instances of default morphology. A crucial example against the generalization that defaults involve the omission of inflection comes from L2 Dutch gender; although we see an alternation between the presence and absence of morphology – alternatively, between an overt morpheme and a zero morpheme – the default is not the morphological zero. Sabourin (2003) reports on the overuse of common gender in neuter contexts, suggesting that common gender acts as a default in determiners and adjectives. The contrast between common and neuter adjectives involves the presence vs. absence of an overt morpheme, as shown in (8–9): *-e* corresponds to common gender, and zero corresponds to neuter gender. These data come from a grammaticality judgement task in which Sabourin’s participants were given sentences like the one in (10), with either the correct or the incorrect adjective.

- 8) een klein-e tafel
 a small-COM table-COM
- 9) een klein kind
 a small-NEUT child-NEUT
- 10) Hij loopt op een gekke/*gek manier.
 he walks in a funny-COM/*funny-NEUT way-COM
 ‘He walks in a funny way.’

Participants were more likely to accept common *-e* in contexts in which it did not belong than to accept a missing *-e* in contexts in which it

should have occurred (e.g. *een klein tafel ‘a small-NEUT table-COM’). Default common *-e* is therefore a crucial example, as it illustrates the inadequacy of an account based solely on the absence of morphology. Put simply, defaults do not equal zeros. Furthermore, these studies on L2 Dutch also find that participants use common gender *de* as a default determiner in place of the neuter *het*, suggesting that the source of the problem may lie in the features associated with these morphemes, rather than in the use vs. non-use of overt morphology.

Crucial to the discussion of representation and computation is the effect of task: the observation of morphological variability in production is, in some ways, inconclusive. As mentioned above, some authors take its occurrence to be indicative of underlying knowledge: variability equals a lack of (syntactic) competence. For others, especially those who argue for the MSIH, variability in production is attributed to problems of lexical access; difficulties arise when the pressure to communicate is strong. For instance, Montrul (2004: 371) concludes that L2 Spanish morphological errors arise because ‘the mature *performance* system is set in a way that becomes inflexible to accommodate structures that differ from those of the L1’ (emphasis added). In principle, then, variability in production may mean either that L2 knowledge is truly non-targetlike, or that it is (more) targetlike but that we cannot gain insight into this knowledge given the nature of the task. If errors are merely a product of performance limitations, then we should be able to observe something closer to the underlying competence by engaging L2ers in tasks that minimize the pressure placed on the learner. As Prévost and White (2000) suggest, this can be done by employing receptive tasks, specifically comprehension and grammaticality judgment tasks.

In addition to the issue of variability across tasks, it is crucial to ask whether comprehension variability, if it exists, is of a similar nature to production variability. For example, do learners employ the same default form(s) that they adopt in production? If they do, this would constitute an argument for an underlying representational issue driving the use of default morphology across methodologies. To date, L2 morphology studies within the generative tradition address this point only indirectly. Franceschina (2002) finds that variability in L2 Spanish gender extends to comprehension but does not identify the types of

errors or the defaults employed, arguing that the existence of morphological variability means that English natives cannot acquire gender as a functional feature in L2. White *et al.* (2004) report variability in the comprehension of Spanish gender; across tasks, L2ers employ masculine defaults. Unexpectedly, participants' overall performance on gender items on the comprehension task is significantly worse than on the production task. White *et al.* note that the use of defaults in comprehension counters Prévost and White (2000)'s suggestion that defaults are confined to production, and poses a challenge to the MSIH. Clearly, the relationship of variability in production to variability in comprehension is an area that requires further exploration.

To summarize, studies of morphological variability in verbal and nominal morphology lead to the following preliminary generalizations:

- Variability does not entail syntactic deficits (i.e. problems with word order and Case).
- Variability typically yields a default – the outcome of systematic substitution errors – that may be either zero or overt.
- Variability extends across proficiency levels.

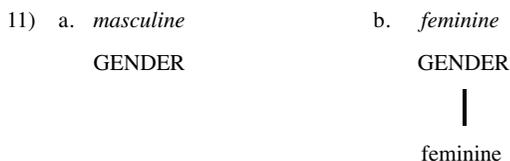
In this article, I propose that two generalizations be added to this list: that variability extends to comprehension, and that variability in comprehension is qualitatively similar to variability in production. These generalizations follow from a set of experiments described in Sections V and VI. These experiments show that the representation of features is crucial to our understanding of morphological variability. The following section briefly outlines one approach to modelling features in morphological theory.

III The representation of features in morphological theory

Under current proposals in morphological theory, features are represented not as unstructured bundles, but are instead modelled in hierarchical feature geometries (e.g. Harley, 1994; Bonet, 1995; Harley and Ritter, 2002). Feature geometries offer the advantage of modelling natural classes such as person, number and gender: feature categories that native speakers possess inherent knowledge of. Natural class nodes are characterized as organizing nodes that dominate daughter nodes – the

features that fill in the members of these natural classes – so that the presence of a node such as ‘feminine’ entails the presence of the dominant, natural-class node of ‘gender’. The feature [feminine] is therefore dependent upon the existence of [gender]; likewise, the feature [plural] is dependent upon the existence of [number].³ (I refer to [feminine] and [plural] as ‘feature values’; these are opposed to the values [masculine] and [singular]). Furthermore, geometries encode markedness via the number of nodes; marked feature values contain additional structure relative to their unmarked counterpart. Harley and Ritter (2002) note that although feature dependencies must be a universal property of language, morphological theory has often failed to attempt to explain them. For the purposes of L2 acquisition, feature geometries provide a theoretical basis for discussing how features are represented in the interlanguage grammar.

For the feature of gender, the dependent features of masculine and feminine are, by assumption, not represented symmetrically. Instead, markedness is encoded by the additional structure accorded to the marked form: feminine. The unmarked value is encoded by under-specification: masculine is encoded as the presence of a bare gender node without any further specification. The representation of masculine and feminine gender in a partial feature geometry is shown in (11).



Markedness relations may be determined independently, rather than through stipulation, as follows. The neutralization of contrast suggests that masculine gender is unmarked in Spanish. The plural of *hermano* ‘brother’ is *hermanos* ‘brothers/siblings’, which can include male and female siblings; the plural of *hermana* ‘sister’, *hermanas*, can only include female siblings. Furthermore, broader syntactic distribution indicates unmarked status; Harris (1991) has argued, on the basis of facts about the Spanish lexicon, that masculine is unmarked – and, by

³ Harley and Ritter (2002) employ the features [class] rather than [gender], and [group] rather than [plural]. The precise terminology is not essential here.

assumption, underspecified – in Spanish. Masculine agreement surfaces in a wider range of syntactic contexts, including those in which no trigger for agreement is present. One example that supports his claim is the use of the preposition *para* ‘for’:

- 12) Tienes demasiados ‘paras’ en este párrafo.
 You-have too-many-MASC paras in this paragraph.
 (Harris, 1991)

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. I further assume, following these authors, that markedness is encoded via underspecification. In addition, I assume that L2 learners – given the availability of evidence from examples parallel to (12) and knowledge of the meaning of words like *hermanos* – similarly come to represent masculine as unmarked and underspecified.

Turning from gender to number, the same logic applies: [number] is an organizing node, with its dependant node, [plural]. Singular number is unmarked and, by assumption, underspecified. We may establish this relationship by examining their syntactic distribution. The question word *quién* ‘who’ is used when it is not known whether the answer is singular or plural. *Quiénes* ‘who-PL’, however, has a restricted distribution: it may only be used in those contexts in which the speaker presupposes that the answer is plural. This contrast is shown in (13).

- 13) a. ¿Quién comió las galletas? (expected answer: singular or plural)
 who ate-3SG the cookies
 b. ¿Quiénes comieron las galletas? (expected answer: plural)
 who-PL ate-3PL the cookies

To explain morphological variability, we require a mechanism to ensure that underspecified forms act as defaults. Such a mechanism is outlined in detail in McCarthy (2007: 62), and shown in (14).

- (14) L2 errors are instances of underspecification, not feature clash.

Combining our representations and the mechanism in (14) – and embedding these assumptions within a framework like Distributed Morphology which allows for the underspecification of vocabulary items relative to syntactic feature specifications (Halle and Marantz, 1993) – a set of predictions regarding substitution errors may be derived. Two pairs of hypothetical errors in clitic agreement are presented in (15) and (16).

Assuming the full specification of features in syntax, feature clash would arise when the syntactic context is (fully) specified as masculine, but a feminine form is inserted, resulting in (15a); an underspecification error would arise when the syntactic context is specified as feminine, but an underspecified masculine form is inserted, resulting in (15b). Similarly, when the syntactic context is (fully) specified as singular, but a plural form is inserted, the result is (16a); an underspecification error as in (16b) would arise when the syntactic context is plural, but an underspecified singular form is inserted. The asymmetrical representation of features allows for the prediction that substitution errors are bidirectional: the (a) examples below trigger feature clash, and are excluded from the set of errors. The (b) examples involve underspecification of the clitic relative to the syntactic context, and are predicted to occur, should errors arise.

- | | | | | |
|-----|----|-----------------------|------------|----------------|
| 15) | a. | Tiene un periódico. | *La | está leyendo. |
| | | Has a newspaper-MASC. | CL-SG-FEM | is reading. |
| | | | | (Target: lo) |
| | b. | Tiene una manzana. | *Lo | está comiendo. |
| | | Has a apple-FEM. | CL-SG-MASC | is eating. |
| | | | | (Target: la) |
| 16) | a. | Tiene una pelota. | *Las | está lanzando. |
| | | Has a ball-FEM. | CL-PL-FEM | is throwing. |
| | | | | (Target: la) |
| | b. | Tiene unos cuadernos. | *Lo | está metiendo |
| | | Has some notebooks. | CL-SG-MASC | is putting |
| | | en su mochila. | | |
| | | in her backpack. | | (Target: los) |

IV Research questions

The research questions to be addressed in the present study are as follows:

- Does morphological variability extend from production to comprehension?
- If it does extend to comprehension, is comprehension variability qualitatively similar to production variability?

If the answer to the first question is no, then a computational account such as the MSIH, which attributes variability to lexical access, will be supported. The second question specifically addresses default morphology across tasks, and asks whether masculine and singular defaults surface

across tasks, as predicted by an account based in the underspecification of features. An affirmative answer to both questions will constitute evidence for a representational account of variability: specifically, an account that capitalizes on the representation of morphological features. In Section VII, I argue that the representational account that best models variability places the source of errors in the morphology, rather than the syntax.

V Method

1 Participants

Participants were 24 speakers of L2 Spanish recruited from Spanish language courses at a major university in Ontario, Canada. All were native English speakers. They had first been exposed to Spanish in their early teens or later, making them post-critical-period learners. Most participants reported having had some exposure to French as a subject in school, but none had participated in French immersion programs, and none reported being a bilingual in French and English. The comprehension task was presented first, followed by the production task.

Proficiency of the L2 participants was measured by combining the scores on:

- a cloze test from the Diploma de Español como Lengua Extranjera (Spanish Embassy, Washington, DC, USA); and
- a multiple choice test from the reading/vocabulary section of the MLA Cooperative Foreign Language Test (Educational Testing Service, Princeton, NJ, USA).

Of these 24, nine scored in the advanced range, and 15 in the intermediate range. I adopted the same proficiency level criteria as those adopted in a previous study of L2 Spanish (White *et al.* 2004). Two participants were excluded due to low proficiency; these participants lacked sufficient vocabulary to complete the elicited production task. Ten native speakers of Spanish also participated, with the purpose of testing whether morphological variability might extend to native grammars.

2 Experiment 1: Elicited production:

The production portion of this study examines agreement in both direct object (DO) clitics and adjectives. DO clitics are inflected for gender

and number agreement with their referent. Sentence (17) is an example of two sentences: the first contains a full NP (*una manzana*), and the second a DO clitic (*la*). Clitics may appear after the participle, as in (17), or before the auxiliary, as previously shown in (15) and (16).

- 17) Tiene una manzana. Está comiéndola.
 Have-3SG an-FEM apple. Is eating-CL-FEM
 'He has an apple. He's eating it.'

The use of a DO clitic vs. a full NP is constrained by discourse factors, in that NPs that are actively in consideration due to recent mention are likely to be realized as clitics. Clitics were elicited by creating contexts in which the NP was recently mentioned (see Procedure).

Adjectives in Spanish also agree in gender and number with the head noun, which may or may not be overt. This task elicited both predicative adjectives (18a) and attributive adjectives (18b), both of which are domains of gender and number agreement.

- 18) a. La camisa es blanca.
 The-FM shirt is white-FEM.
 b. Lleva una camisa blanca.
 She's-wearing a-FEM shirt white-FEM.

a Materials: Test materials consisted of 20 colour photographs. These photographs pictured an agent acting upon an object; the latter was intended to be realized as a DO clitic. For example, one picture displayed a boy holding an apple up to his mouth, about to bite into it. Another displayed two girls, the first holding a pair of used textbooks, the second holding money, apparently about to buy the books. The objects were chosen with the goal of eliciting 10 masculine noun phrases and 10 feminine ones. In addition, 10 of the pictures were intended to elicit singulars, and 10 to elicit plurals. These items are listed in Table 1.

b Procedure: Participants were interviewed individually by a native speaker of Spanish. Each participant was shown all 20 pictures, one at a time, in random order. The experimenter initially asked a question designed to elicit the target noun phrase, as in (19). The participant would typically respond by naming the item along with a determiner, as in (20).

Table 1 Elicited production task items by gender and number

	Masculine	Feminine
Singular	el arete 'the earring' el café 'the coffee' el pájaro 'the bird' el libro 'the book' el periódico 'the newspaper'	la manzana 'the apple' la carta 'the letter' la taza 'the cup' la ventana 'the window' la pelota 'the ball'
Plural	los cuadernos 'the notebooks' los libros 'the books' los zapatos 'the shoes' los pantalones 'the pants' los lápices 'the pencils'	las tijeras 'the scissors' las botas 'the boots' las hojas 'the leaves' las revistas 'the magazines' las camisetas 'the shirts'

19) ¿Qué tiene el chico en la mano?
What has the boy in the hand?

20) Una manzana.
An-FEM apple.

Immediately after naming the object, the experimenter asked a question intended to elicit a clitic, as in (21) and (22).

21) ¿Qué va a hacer con la manzana?
What going-to do with the-FEM-SG apple?
'What is he going to do with the apple?'

22) Va a comerla.
Going-to eat-CL-FEM-SG.
'He is going to eat it.'

Following the description of the action, the experimenter asked questions about the colour of the object, and of other objects in the photo. These questions were designed to elicit adjectives as a domain of gender and number agreement. To serve as distracters from the test questions, the interviewer then asked a variety of other questions; for example:

- Where is the girl?
- What else do you see in the picture?
- What clothes is she wearing?, etc.

This procedure was repeated for all 20 pictures. Interviews lasted between 10 and 30 minutes.

c Data analysis: All interviews were transcribed by a near-native speaker of Spanish. Productions of clitics and adjectives were coded twice:

once for accuracy of gender agreement, and once for accuracy of number agreement. It was also noted whether or not the participant had previously named the target object with the correct gender on the determiner; only data for items that the participant was able to name and provide the gender of are reported. I revisit the issue of knowledge of masculine and feminine gender in Section VII, and discuss its impact on accuracy.

3 *Experiment 2: Comprehension*

a Materials: Test materials, adapted from White *et al.* (2004), consisted of a booklet containing a story about two characters going on vacation. The story contained 48 test sentences. Sixteen of these tested gender (8 masculine clitics, 8 feminine clitics) and 16 tested number (8 singular clitics, 8 plural clitics). The remaining 16 were distracters that did not contain clitics. After participants read a sentence containing a clitic, they circled the picture, out of three possible choices, that correctly corresponded to the story. One picture corresponded in gender and number to the clitic. Another picture disagreed in the category of the test item, holding the other category constant. The third picture was a foil: half of the foils disagreed in one category, and half disagreed in two categories. Each picture was designed to be equally plausible in the context of the story. See Figure 1. Example (22) is a sample gender item.

- 22) Paco quiere llevar algunas cosas que acaba de comprar pero no encuentra nada. Paco dice: 'Acabo de comprarlo: ¿dónde está?'

Paco wants to bring some things that he just bought, but he can't find anything. Paco says, 'I just bought it-MASC: Where is it?'

The clitic in (22) is masculine and singular: it agrees with *el cinturón*, the target answer. If the participant chose *la camisa* ('the-FEM shirt'), this would constitute an (underspecification) error, as the participant

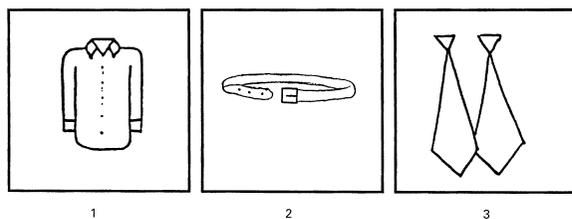


Figure 1 Comprehension task: sample test item containing *camisa* (shirt-FEM), *cinturón* (belt-MASC) and *corbatas* (ties-FEM)

selected a feminine syntactic context to correspond to a masculine clitic. *Las corbatas* ('the-FEM ties') is the foil: it disagrees in both gender and number. Neither the name nor the gender of the items in the pictures appeared in the story. This ensured that participants relied on their knowledge of gender and number agreement in order to select the correct item. Knowledge of the name and gender of the test items was independently established by the vocabulary test, described below.

b Vocabulary test: The vocabulary test consisted of 48 pictures, all of which appeared in the picture selection task. For each item, participants wrote the name of the object and circled the article that corresponded in gender. The vocabulary test was administered before the comprehension task. In some cases, alternate names were given for test items (e.g. *la bola* for *la pelota*: both mean ball, although *pelota* is more common); these were included as correct responses. Where alternate names with opposite gender from the intended target were given (e.g. *la taza* 'the cup' for *el vaso* 'the glass'), items were discarded from the analysis of the comprehension task results, as were items that were left blank.

VI Results

Accuracy rates were analysed statistically with a series of two-way mixed ANOVAs within the General Linear Model of SPSS (Statistical Package for the Social Sciences, version 16 for Macintosh). For each analysis, group is the between-subjects factor (intermediate, advanced, native) and feature value is the within-subjects factor (masculine vs. feminine gender; singular vs. plural number). Main effects of group and feature value, as well as interaction effects, are presented below. *Post hoc* Scheffé tests were performed for all significant between-group differences.

1 Experiment 1: Production

a Clitics: The elicited production task elicited a total of 390 tokens of gender agreement, and 405 tokens of number agreement⁴ in clitics from the L2 participants. Of the 390 tokens of gender agreement, 277 are included in the present analysis; these are the clitics that correspond

⁴ These numbers are not identical because of the participants' tendency to reduce vowels. Productions of phonetic [lə] or [ləs] were coded for number, but not for gender.

to an object that the participant was able to name along with an agreeing determiner. This exclusion of a large number of test items resulted in the exclusion of three intermediate-level participants for gender in clitics; these participants produced fewer than three usable test items for one or both genders. These three participants were included in the analysis of number. Table 2 summarizes the mean accuracy rates for gender and number agreement in clitics. Native speakers (NSs) show no evidence for variability in clitic agreement whatsoever.

Beginning with the analysis of gender, the main effect of group is significant ($F = 6.94; p < .01$). The *post hoc* analysis indicates that this effect is due to the difference between the intermediate and NS groups. The main effect of gender feature value is significant ($F = 7.91; p < .01$), with accuracy greater for masculine contexts than feminine contexts, suggesting the use of default morphology. The interaction between gender and group is not significant ($F = 2.38; p < .2$). A paired-samples *t*-test indicates a feature asymmetry for the advanced ($t = 2.37; p < .05$) group, with accuracy higher in masculine contexts than feminine ones; the intermediate group's significance on this measure fails to reach significance but is indicative of a trend ($t = 2.21; p < .06$). For number, the main effect of group is not significant ($F = .90; p < .5$), nor is the main effect of feature value ($F = 1.79; p < .2$). Although the outcome is not significant, all of the seven errors of number involved the substitution of singular for plural. The interaction between group and number feature value is not significant ($F = .90; p < .5$).

b Adjectives: The elicited production task yielded a total of 378 tokens of gender agreement from the L2ers, 273 of which were retained for analysis as the participant had successfully named the object along

Table 2 Mean percent accuracy in clitic agreement on the production task, by group and feature

	Intermediate	Advanced	Natives
Gender (all)	82.3	92.0	100.0
Masculine	94.8	100.0	100.0
Feminine	69.9	84.0	100.0
Number (all)	96.8	98.9	100.0
Singular	100.0	100.0	100.0
Plural	93.7	97.9	100.0

with a correctly inflected determiner. The task elicited 488 tokens of number agreement in adjectives from the L2 participants. One of the advanced-level participants was excluded from the gender analysis for producing fewer than three usable tokens in one condition; one native speaker was excluded from the number analysis for failing to produce sufficient tokens of plural adjectives. Table 3 summarizes the accuracy rates for gender and number agreement in adjectives by proficiency group and feature. Once again, we find variability in the production of agreement; gender agreement similarly lags behind number agreement. However, the accuracy rate with number in adjectives is well below the rate in clitics, with neared 100% across the board.

The main effect of group is significant for gender ($F = 12.46$; $p < .001$); according to the *post hoc* test, this effect arises from the contrast between the intermediate and both other groups. The main effect of gender feature value is also significant ($F = 11.56$; $p < .003$), with masculine agreement significantly more accurate than feminine agreement. The interaction between group and feature value is also significant ($F = 6.52$; $p < .005$), indicating the effect of feature is not equal across groups. There is a large spread between the two values for the intermediate group (62.7% for feminine vs. 96.7% for masculine); the difference in accuracy between masculine and feminine contexts is significant for the intermediate group ($t = 3.95$; $p < .002$). The advanced group shows no significant difference between the two feature values ($t = 1.77$; $p < .2$) and shows a spread of only about 10 points (86.9% for feminine vs. 95.7% for masculine).

Turning to number in adjectives, the main effect of group is significant ($F = 3.47$; $p < .05$); the *post hoc* test indicates that this effect is due to the difference between the intermediate and advanced groups. The main

Table 3 Mean percent accuracy in adjective agreement on the production task, by group and feature

	Intermediate	Advanced	Natives
Gender (all)	79.7	90.7	100.0
Masculine	96.7	95.7	100.0
Feminine	62.7	86.9	100.0
Number (all)	84.0	91.9	100.0
Singular	99.6	100.0	100.0
Plural	68.2	83.9	100.0

effect of number feature value is also significant ($F = 9.73$; $p < .005$), with singular agreement more accurate than plural. This result follows from the use of singular agreement as a default in plural contexts. The interaction between group and number feature value is significant ($F = 3.50$; $p < .05$). Unlike clitics, for adjectives both the intermediate and advanced groups show a spread of over 15 points between accuracy rates for masculine and feminine contexts. For the intermediate group, the difference in accuracy between masculine and feminine is significant ($t = 3.33$; $p < .006$); for the advanced group it is not significant ($t = 1.86$; $p = .10$). A closer look within the advanced group, however, reveals that plural number is variable for some, as three of nine advanced participants used singular agreement in plural contexts. Apparently, plural number remains problematic, even for some speakers at the advanced proficiency level.⁵

2 Experiment 2: Comprehension

The comprehension task produced responses to 380 clitics. The accuracy rates for category and feature by group are presented in Table 4. Beginning with the results of the ANOVAs, the main effect of group is significant for gender ($F = 6.191$; $p < .006$); the *post hoc* test attributes this effect to the intermediate group's performance vs. the two other groups. Importantly, the intermediate group's performance is significantly worse than the native group, indicating variability in comprehension for gender. For number, the main effect of group is not significant ($F = 2.12$; $p < .2$), providing no evidence for variability in the comprehension of number. Turning to feature value, the main effect is significant for gender ($F = 6.96$; $p < .02$), with the comprehension of feminine clitics in masculine contexts more accurate than the reverse. Recall that the production task measured the choice of clitic morphology in response to a given (syntactic) context; here, we measure the reverse: the choice of syntactic context relative to clitic morphology. The relative accuracies for masculine and feminine therefore appear reversed. In comprehension, the

⁵ As one *Second Language Research* reviewer points out, it may be the case that the input the L2ers receive is variable, in that native speakers variably delete plural marking on adjectives (Cedergren, 1973; Poplack, 1980). The L2ers, then, may be faced with variable input in adjectives, which might explain why their accuracy plural adjective agreement is suppressed. Nevertheless, I did not find any evidence for the deletion of plural *-s* among the natives in their performance on the production task.

Table 4 Mean percent accuracy on the comprehension task, by group and feature

	Intermediate L2	Advanced L2	Natives
Gender (All)	79.7	97.3	99.0
Masculine	68.4	96.3	98.0
Feminine	90.9	98.4	100.0
Number (All)	97.5	96.5	100.0
Singular	98.3	95.8	100.0
Plural	96.7	97.2	100.0

lower accuracy of masculine gender indicates that a masculine clitic is more likely to be interpreted as referring to a feminine object than the reverse, indicating the use of a masculine clitic as a default in comprehension. A significant interaction between group and gender feature is also found ($F = 4.78$; $p < .02$), with the accuracy of the intermediate group suppressed for only one feature value: masculine in feminine contexts, at 68.4%. The intermediate group's mean score for feminine clitics in masculine contexts is over 20 points higher, at 90.9%; the difference between accuracy scores is significant ($t = 3.31$; $p < .01$). To compare, the gap between feature values for the advanced group is only 2% (96.3 for masculine clitics vs. 98.4 for feminine clitics) and is not significant ($t = .50$; $p < .7$). For number, the main effect of feature value is not significant ($F = .012$; $p < 1$), as both singular and plural scores approach ceiling. Of the 11 errors in the comprehension of number, five involved singular clitics and six plural. Again, the high rate of accuracy in number indicates near-ceiling performance on the comprehension task.

A comparison of accuracy rates for clitics across tasks reveals that L2ers are more accurate on items testing for number – which nears ceiling and never drops below 95% – than for gender, which shows a great deal of variability in both comprehension and production. For gender, the effect of task differs by proficiency group. For intermediate participants, accuracy shows a slight drop in comprehension (from 82.3 for production to 79.7 for comprehension), but for the advanced group, accuracy is higher in comprehension (from 92.0 for production to 97.3 for comprehension).

VI Discussion

In light of the data from these two experiments, we may answer the two research questions in the affirmative: the difference in accuracy

between NSs and L2ers indicates that morphological variability does extend to comprehension, and given the use of masculine defaults across tasks, variability is qualitatively similar across comprehension and production. The theoretical significance of these observations, and their relationship to our current understanding of morphological variability, is discussed below.

First, the results of the elicited production experiment confirms that variability is a persistent phenomenon found at even very high levels of proficiency. This observation adds to the body of research on the persistence of L2 morphological variability, including Lardiere (1998a; 1998b), Hawkins (2000) and White (2003) for English, and Franceschina (2001) for Spanish. Of the nine advanced speakers who participated in the elicited production task, only two showed 100% accuracy in agreement across the board.

The following are examples of gender errors in clitics (23), and adjectives (24, 25) made by speakers at advanced levels of proficiency. The examples include dialogue between the experimenter (E) and the participant (P) in which the participant correctly names the item and provides the correct determiner.

Participant 27, Advanced L2 Spanish

- 23) E: ¿Qué está tocando el chico con las manos?
'What is the boy touching with his hands?'
- P: La ventana.
'The-FEM window.'
- E: ¿Qué es lo que quiere hacer con la ventana?
'What is it that he wants to do with the window?'
- P: Lo quiere cerrar.
CL-MASC-SG wants to-close
'He wants to close it.'

Participant 16, Advanced L2 Spanish

- 24) P: Tiene tijeras, y está poniendo las tijeras en la mochila. [...]
'She has scissors, and she's putting the scissors in the-FEM backpack.'
- E: ¿De qué color es la mochila?
'What colour is the backpack?'
- P: La mochila es negro con un poco de blanco.
The-FEM backpack is black-MASC with a little white.
'The backpack is black with a little white.'

Participant 33, Advanced L2 Spanish

- 25) *P*: Está poniendo las zapatillas.
 ‘She’s putting the-FEM-PL sneakers.’
E: ¿De qué color son las zapatillas?
 ‘What colour are the sneakers?’
P: Blancos.
 White-MASC-PL

Gender errors in (23–25) contrast with the appropriate use of gender by the same participants in (26–28); seen together, they demonstrate morphological variability.

Participant 27, Advanced L2 Spanish

- 26) *E*: Este chico, ¿qué tiene en la mano?
 ‘This boy, what does he have in his hand?’
P: Una manzana.
 An-FEM apple
E: ¿Qué está haciendo con la manzana?
 ‘What is he doing with the apple?’
P: La va a comer.
 CL-FEM going-to eat
 ‘He’s going to eat it.’

Participant 16, Advanced L2 Spanish:

- 27) *E*: ¿De qué color es la mesa?
 ‘What colour is the table?’
P: La mesa es blanca.
 The-FEM table is white-FEM
 ‘The table is white.’

Participant 33, Advanced L2 Spanish:

- 28) *E*: ¿Cómo es la ropa de las chicas?
 ‘What are the girls’ clothes like?’
P: Tiene pantalones blancos, camiseta negra.
 She-has pants-MASC white-MASC-PL, t-shirt-FEM black-FEM
 ‘She has white pants, black shirt.’

As noted in Section VI, number agreement also remains somewhat variable for advanced speakers, particularly in the domain of adjectives; predicative adjectives appear particularly difficult. In (29), the participant uses singular agreement in a plural context. Here, the use of a plural copula *son* makes the syntactic context unambiguously plural. At the same time, the participant uses masculine gender in a feminine context.

Participant 16, Advanced L2 Spanish

- 29) E: ¿Y te gustan sus medias, o vas a criticarlas?
 ‘And do you like her socks, or are you going to criticize them?’
- P: No no no. Son gris y blanco.
 No no no. They-are gray-SG and white-SG
 ‘No, no, no. They’re gray and white.’

Second, in addition to its persistence, variability is systematic rather than random. Masculine inflection occurs in feminine contexts, whereas the reverse pattern is rarely attested. Singular inflection occurs in plural contexts (albeit only in production), whereas the reverse pattern is rarely attested. In some domains (e.g. singular number agreement) no evidence of variability is found, making it systematically targetlike. These results can be interpreted as evidence for the emergence of underspecified default morphology (for a more detailed discussion of feature representations, underspecification and markedness, see McCarthy, 2007).

Third, and most crucially, morphological variability extends to comprehension. The intermediate group, in particular, was significantly worse than both other groups in the comprehension of gender overall. This observation is particularly problematic for the MSIH, which holds variability to be due to production problems. Without denying the role of such factors in production, this explanation cannot be maintained in light of the comprehension data.

Fourth, comprehension variability is qualitatively similar to production variability. Across both comprehension and production, L2ers employed masculine defaults. This suggests a representational asymmetry between masculine gender on one hand, and feminine gender on the other. For number, singular was also employed as a default in the production of both clitics and adjectives; the comprehension of number, however, proved unproblematic.

A closer look at this feature asymmetry reveals an additional effect of knowledge of gender. Up to this point, I have only reported on the responses in which the participant had demonstrated knowledge of gender by naming the pictured object along with a determiner. However, it is worth noting the behaviour on items in which knowledge of gender was not demonstrated; that is, the cases in which the participant was not able to name the item in the picture, or named it but could not produce a correctly agreeing determiner. I divided the entire dataset of produced

clitics into two categories: those in which knowledge of the object's gender was demonstrated, and those in which it was not; each error was then coded for knowledge (demonstrated or not demonstrated) and target-gender feature value. The effect of knowledge – whether demonstrated or not demonstrated – plays out differently depending on feature value. Table 5 breaks down clitic errors by target-gender feature value and knowledge of gender. Of the 49 errors in the production of feminine clitics, 33 (or approximately two-thirds) were cases in which the participant had previously named the object along with its agreeing determiner, as in (23). Of the 16 errors in masculine gender, only 6 – fewer than half – occurred when the participant had previously named the object. The contingency between feature value and knowledge of gender is significant, as measured by a chi-square test ($\chi^2 = 26.82$, $df = 1$, $p < .0001$). This pattern suggests that when participants know an object to be masculine, they generally do not use the inappropriately valued clitic; the same cannot be said for feminine gender. The contrasting behaviour of feature values suggests a fundamental asymmetry between the two feature values in opposition: errors in masculine agreement surface when the participant does not know the gender of the target item (and is forced to guess), whereas errors in feminine agreement surface both when gender is known and when it is not known. From the persistent variability in gender usage in contexts where gender is known – specifically when feminine gender is at stake – it is apparently not the case that correct encoding of gender automatically equals the correct use of agreement (see Carroll, 1989). The suppliance of agreement in feminine contexts remains problematic despite the knowledge of feminine gender. When gender of an object is known, masculine clitic agreement appears to 'come for free'; feminine agreement, however, apparently does not.

Table 5 Distribution of errors in clitics for all L2 participants on the production task: target gender feature value vs. knowledge (whether demonstrated or not demonstrated)

	Masculine	Feminine	Totals
No knowledge	10	16	26
Knowledge	6	33	39
Totals	16	49	65

Returning to the issue of representation vs. computation, while these experimental data suggest a representational motivation for variability, current representational accounts (e.g. Hawkins and Chan, 1997) are not fully consistent with syntactic patterns reported in previous studies. The goal of this article, however, is not to argue that syntactic deficits cannot exist in the L2 grammar; I simply take the position, as noted in Section I, that syntactic deficits are not entailed by morphological variability. In a previous study of L2 Spanish word order, White *et al.* (2004) report 100% accuracy in noun–adjective word order among intermediate and advanced learners of L2 Spanish,⁶ despite the occurrence of variability in gender agreement on adjectives. White *et al.* conclude that the features and feature strength triggering N-raising to NUM are not impaired, indicating that there is no underlying syntactic impairment, if one takes accuracy in word order as indicative of underlying syntactic competence. White *et al.*'s claim that syntax is unimpaired – combined with the present study's observation of variability in comprehension – suggests that any representational 'deficits' lie in the morphology, rather than the syntax.

With the assumption of asymmetrical representations in the morphology in place, there are (at least) two ways that these representations can be understood to contribute to variability. As a first possibility, we may propose that L2ers' representations are completely native-like: we may assign them the asymmetrical representations like the ones in (11a–b). The problem, then, would lie simply in overusing the representationally simpler form over the other: an issue of 'performance'. Alternatively, we may assume that L2ers do not necessarily represent features in a native-like way; instead, their feature representations are impoverished in a way that is non-nativelike: an issue of 'competence'. These two explanations make different predictions, as I show below.

The first account, which assumes native-like morphology, would encounter the same problem faced by the MSIH. We might expect morphological errors to occur more frequently in production, assuming that production is more taxing. As shown in Section VI, this prediction is not supported for the intermediate group. Furthermore, we might also expect L2ers' accuracy to increase generally, as performance comes to

⁶ The proficiency level of White *et al.*'s speakers and those whose data is reported here is comparable, since our task employed the same proficiency test and same group criteria. Furthermore, in the production task I found no instances of *Adjective Noun word order in the entire dataset.

reflect competence more closely. Errors would not be predicted to change in any particular way as proficiency increases. Although we do find effects of group for all three measures, the clitic production data suggest that the two L2 groups may be treating features differently, as I discuss below.

Under the second account, however, we may assume that L2 speakers' representations are, in some cases and particularly at lower proficiency levels, deficient; therefore, we correctly predict variability across tasks. Some L2ers may lack the dependent feature [feminine], whereas other, presumably higher-proficiency ones, do represent it. Prior to the acquisition of [feminine], no feature asymmetry would exist. At this early stage, errors are predicted to be bidirectional: underspecification errors are predicted in both directions, as neither feature is specified. A subset of the intermediate-level participants could reasonably be placed in this stage: five of 15 make errors in clitic agreement that should be ruled out by the ban on feature clash as stated in (14), while the remaining 10 use systematic defaults. This contrasts with the advanced group, where none of the nine participants produced an error of feature clash in clitics. As the marked feature is acquired, a feature asymmetry emerges; errors become unidirectional, and errors in the unmarked masculine context are excluded under (14). Tentatively, we may view the clitic results as evidence for the second account: advanced participants reached 100% accuracy in masculine clitic agreement, whereas the intermediate group remained somewhat variable on both features; recall that the intermediate group's feature asymmetry failed to reach significance. Adjective agreement gives a different view, however. Neither group makes many errors in masculine agreement at all, yielding systematic substitutions of masculine for feminine for all proficiency groups; potentially this follows from the use of masculine adjectives as citation forms (e.g. *rojo* 'red-MASC' means 'red'). While the case for asymmetrical representations in L2 morphology appears strongly supported based on the data reported here, further research is needed to assess whether the suggestion of morphological development is on the right track.

VIII Conclusions

Morphological variability is, at least in part, a representational issue, and does not derive strictly from production-based limitations (see Prévost and White, 2000). In Lardiere's (2005) terms, the acquisition of

feature representations can be considered a part of *morphological competence*. However, morphological knowledge need not be considered a different kind of knowledge than syntactic or phonological knowledge. All of these areas are domains that consist of organized, structured representations, and L2 research can aim to discover how the competence that consists of these representations is acquired.

Of course, the role of performance cannot be ruled out entirely. Crucially, under the current proposal, performance is not thought to be limited to production, as it appears to be under the MSIH. Linguistic competence is accessible only indirectly (see White, 2003: 17), and performance may have a greater or lesser effect on the data we access depending on the task. Correspondingly, we may see more variability in one task as opposed to another; one example of this is the climb in accuracy from production to comprehension among the advanced group. This proposal predicts that variability will surface in qualitatively similar ways across comprehension and production; furthermore, it does not entail the presence of syntactic deficits, a point of controversy in recent L2 theory. Finally, the adoption of hierarchical features provides a basis for modelling morphological development during the course of L2 acquisition.

Acknowledgements

This research project was supported by the following funding sources: SSHRCC 410–2004–1774 (to Lydia White), and FQRSC 2001-ER-66973 and 2006-SE-103710 (to Lydia White *et al.*), as well as the George Mason University College of Humanities and Social Sciences. The author would like to acknowledge the following for their contributions to this work: Lydia White, Glyne Piggott, Andrea Gualmini, Susana Bejar, Joyce Bruhn de Garavito, Ana Faure, Tel Monks, Roumyana Slabakova, Elena Valenzuela, and the comments of three anonymous reviewers. Any errors and inconsistencies remain my own.

IX References

- Bonet, E.** 1995: Feature structure of romance clitics. *Natural Language and Linguistic Theory* 13, 607–47.
- Carroll, S.** 1989: Second-language acquisition and the computational paradigm. *Language Learning* 39, 535–94.

- Cedergren, H.** 1973: Interplay of social and linguistic factors in Panama. Unpublished PhD dissertation, Cornell University.
- Clahsen, H.** 1988: Parameterized grammatical theory and language acquisition: a study of the acquisition of verb placement and inflection by children and adults. In Flynn, S. and O'Neil, W., editors, *Linguistic theory in second language acquisition*. Dordrecht: Kluwer, 47–75.
- Franceschina, F.** 2001: Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research* 17, 213–47.
- 2002: Case and ϕ -feature agreement in advanced L2 Spanish grammars. In Foster-Cohen, S.H., Ruthenberg, T. and Poschen, M.L., editors, *EUROSLA Yearbook 2*, Amsterdam: John Benjamins, 71–86.
- Halle, M. and Marantz, A.** 1993: Distributed morphology and the pieces of inflection. In Hale, K. and Keyser, S.J., editors, *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge MA: MIT Press, 111–76.
- Harley, H.** 1994: Hug a tree: deriving the morphosyntactic feature geometry. In Carnie, A., Harley, H. and Bures, T., editors, *MIT Working Papers in Linguistics* 21, 289–320.
- Harley, H. and Ritter, E.** 2002: Person and number in pronouns: a feature-geometric analysis. *Language* 78, 482–526.
- Harris, J.W.** 1991: The exponence of gender in Spanish. *Linguistic Inquiry* 22, 27–62.
- Hawkins, R.** 2000: Persistent selective fossilisation in second language acquisition and the optimal design of the language faculty. *Essex Research Reports in Linguistics* 34, 75–90.
- Hawkins, R. and Chan, C.Y.-H.** 1997: The partial availability of universal grammar in second language acquisition: the 'Failed Functional Features Hypothesis'. *Second Language Research* 13, 187–226.
- Haznedar, B. and Schwartz, B.** 1997: Are there optional infinitives in child L2 acquisition? In Hughes, E. Hughes, M. and Greenhill, A., editors, *Proceedings of the 21st Annual Boston University Conference on Language Development (BUCLD)*. Somerville, MA: Cascadilla Press, 293–306.
- Lardiere, D.** 1998a: Case and tense in the 'fossilized' steady state. *Second Language Research* 14, 1–26.
- 1998b: Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research* 14, 359–75.
- 2005: On morphological competence. In Dekydtspotter, L., Sprouse, R. A. and Liljestr and, A., editors, *Proceedings of the 7th Generative Approaches to Second Language Acquisition (GASLA 2004)*. Somerville, MA: Cascadilla Proceedings Project, 178–92.
- McCarthy, C.** 2007: Morphological variability in second language Spanish. Unpublished PhD dissertation, McGill University, Montreal.

- Meisel, J.** 1991: Principles of Universal Grammar and strategies of language learning: some similarities and differences between first and second language acquisition. In Eubank, L., editor, *Point counterpoint: Universal Grammar in the second language*. Amsterdam: John Benjamins, 324–76.
- Mezzano, G.** 2003: The development of Spanish verbal inflection in early stages of L2 acquisition. Undergraduate honors thesis, University of Illinois at Urbana-Champaign.
- Montrul, S.A.** 2004: *The acquisition of Spanish: morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition*. Amsterdam: John Benjamins.
- Paradis, J.** 2005: Grammatical morphology in children learning English as a second language: implications of similarities with Specific Language Impairment. *Language, Speech and Hearing Services in Schools* 36, 172–87.
- Poeppl, D.** and **Wexler, K.** 1993: The full competence hypothesis of clause structure in early German. *Language* 69, 1–33.
- Poplack, S.** 1980: The notion of the plural in Puerto Rican Spanish: competing constraints on /s/ deletion. In Labov, W., editor, *Locating language in time and space*. New York: Academic Press, 55–68.
- Prévost, P.** and **White, L.** 2000: Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16, 103–33.
- Sabourin, L.** 2003: Grammatical gender and second language processing: an ERP study. Unpublished PhD dissertation, University of Groningen.
- Vainikka, A.** and **Young-Scholten, M.** 1994: Direct access to X'-theory: evidence from Korean and Turkish adults learning German. In Hoekstra, T. and Schwartz, B.D., editors, *Language acquisition studies in generative grammar*. Amsterdam: John Benjamins, 265–316.
- 1996: Gradual development of L2 phrase structure. *Second Language Research* 12, 7–39.
- White, L.** 2003: *Second language acquisition and Universal Grammar: from initial to steady state*. Cambridge: Cambridge University Press.
- 2007: Some puzzling features of L2 features. In Licerias, J., Zobl, H. and Goodluck, H. editors, *The role of formal features in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum, 305–30.
- White, L., Valenzuela, E., Kozłowska-Macgregor, M.** and **Leung, Y.-K.I.** 2004: Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25, 105–33.