SSLA, **31**, 1–29. Printed in the United States of America.
doi: 10.1017/S0272263109090019

WORD CLASS DISTINCTIONS IN SECOND LANGUAGE ACQUISITION

An Experimental Study of L2 Spanish

Eve Zyzik
University of California, Santa Cruz

Clara Azevedo
Michigan State University

Although the problem of word class has been explored in numerous first language studies, relatively little is known about this process in SLA. The present study measures second language (L2) learners' knowledge of word class distinctions (e.g., noun vs. adjective) in a variety of syntactic contexts. English-speaking learners of Spanish from third-semester and third-year courses (N = 240) completed a receptive task that presented contrasting forms belonging to the same word family (e.g., *feliz* "happy" and *felicidad* "happiness"). The results indicate that learners from both groups are often unable to distinguish among word classes. In particular, learners have significant difficulty in discriminating between adjectives and nouns. Although ambiguous surface morphology contributes to word class confusions, the results suggest that L2 learners do not always recognize derivational suffixes that clearly mark word class. These difficulties are interpreted as stemming from weak syntactic morphological knowledge as well as incomplete knowledge of L2 distributional regularities.

We would like to extend our thanks to the many individuals who contributed to this research project, including Charlene Polio, Zach Hambrick, the instructors of the Spanish classes that took part in the study (Nicole Metcalf and Dora Martín-Cabrera), and the students at Michigan State University who volunteered to participate. We are grateful to Regina Morin for providing us with the mean scores from her 2006 paper. Finally, we would like to thank the anonymous *SSLA* reviewers for their useful feedback. All remaining errors are ours.

Address correspondence to: Eve Zyzik, Humanities 131, University of California, 1156 High Street, Santa Cruz, CA 95064; e-mail: ezyzik@ucsc.edu.

© 2009 Cambridge University Press 0272-2631/09 $15.00

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Linguistic descriptions of all languages presuppose the division of words into distinct classes or lexical categories (e.g., nouns, verbs, adjectives, adverbs). Although there is disagreement with respect to the inventory of lexical categories found crosslinguistically and to the criteria used for classifying words into those categories (cf. Baker, 2003), the discovery of such distinctions is a key aspect of learning any given language and is central to the productivity of grammar (Maratsos, 1988). In simple terms, the ability to distinguish among word classes allows the speaker to use particular forms in certain grammatical frames; for example, native speakers of Spanish implicitly know that *triste* "sad" can only be used in an adjective slot (*un caso triste* "a sad case"), whereas the related nominal form *tristeza* "sadness" has a distinct distribution *(no puede superar la tristeza* "he can't get over the sadness"). These forms, although semantically related, are not interchangeable: Using the appropriate form in a given syntactic context is fundamental for producing grammatical sentences in the target language.

Although the problem of how word classes are acquired has been explored theoretically and empirically in the context of first language (L1) acquisition (e.g., Braine, 1987; Gerken, 2001; Gerken, Wilson, & Lewis, 2005; Golinkoff, Hirsh-Pasek, & Schweisguth, 2001; Mintz, 2006; Mintz, Newport, & Bever, 2002), very little is known about this process in SLA. Adult second language (L2) learners have a highly developed system of word classes in their native language—but the question of how much of this knowledge transfers to the L2 remains to be addressed. Whereas some positive transfer is likely, much of an individual's knowledge of word classes is language-specific, given the importance of distributional information in assigning word class (Ellis, 2001; Redington, Chater, & Finch, 1998). We will explore this issue here; however, a starting point for our research is the word class errors that have been documented in a variety of L2 contexts (Bardovi-Harlig & Bofman, 1989; Hakuta, 1987; Lafford & Collentine, 1987, 1989; Odlin & Natalicio, 1982). For example, Whitley (2004) demonstrated a high rate of L2 learners' lexical errors in Spanish due to word class confusions, as illustrated in (1) and (2). The correct form is provided in parentheses.

(1) *la música ruido* (Target: *la música ruidosa*)
"the noise music"

(2) *sin feliz* (Target: *sin felicidad*)
"without happy" (Whitley, p. 166)

The key question is thus why L2 learners use words in inappropriate grammatical frames, despite the fact that, in many cases, such usage would constitute a word class violation in the learners' L1 as well. Whitley suggested that such word class errors are the result of a production strategy: Learners supply the form with which they are familiar in order to fill a lexical gap at the moment of communication. However, an additional complication may be present if

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

learners do not recognize the word class in their own output. As an example, the learner in (1) may have supplied the form *ruido* "noise" based on (a) the misclassification of the word (i.e., the learner thinks *ruido* "noise" could be an adjective) or (b) the failure to recognize the syntactic context as requiring a particular word class (i.e., the learner thinks a noun can appear in this frame). The problem is complex and unlikely to be clarified with production data alone.

The present study reports on the findings of an experiment designed to test learners' knowledge of word classes in their L2 in a variety of syntactic contexts. Specifically, we investigated learners' ability to distinguish between words belonging to the same family when choosing from a set of options. The focus is on receptive knowledge of word classes because we hypothesized that the problem is not solely at the level of production. If errors like (1) and (2) reflect some sort of performance problem or momentary lexical gap, then learners should be able to distinguish between pairs of words like *feliz* "happy" and *felicidad* "happiness" in a receptive task (i.e., when both forms are presented to them as options). On the other hand, if an error is observed in both production and comprehension (i.e., there are across-the-board effects), the problem may involve a competence deficit.[1] After presenting the problem of acquiring word classes in greater detail, we will address the crosslinguistic variation in morphological cues to word classes and present previous findings from the L2 literature.

THE ACQUISITION OF WORD CLASSES

Schmidt (2001), in a discussion of the concept of attention to input in SLA, mentioned the problem of word class:

> The only obligatory consequence of attention to input is that learners become aware of sounds, words (recognizable sequences of sounds associated with meanings), and sequences of words. These words are examples of lexical categories such as noun, adverb, and so on, but the input does not come labeled that way (p. 31).

Learners must parse and syntactically analyze utterances in order to come to know that particular words are exemplars of a given lexical category, although they may not know the grammatical labels or have the metalinguistic knowledge to explain their usage. Research indicates that even native speakers often lack this kind of metalinguistic knowledge (Alderson, Clapham, & Steel, 1997). Schmidt mentioned several possibilities of how lexical categories might become established in a L2: through transfer from the L1, through implicit analysis of the L2 input, or through explicit mechanisms such as instruction and hypothesis testing. Schmidt did not comment further on these proposals but emphasized the importance of the problem: "The way in which learners acquire knowledge of lexical categories [constructions, and rules] is a central issue in SLA" (p. 32).

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Cues to Assigning Word Class

Because the acquisition of lexical categories has been largely overlooked in L2 contexts, it is helpful to consider how L1 researchers have approached this issue (see Labelle, 2005, for a comprehensive review). Redington and Chater (1998) explained the twofold nature of the problem for children learning their mother tongue: (a) discovering that there are different categories of words and (b) learning which words are members of which category. A number of theoretical models have been proposed to account for children's initial entrance into the system of word classes, including semantic bootstrapping (Grimshaw, 1981; Pinker, 1984) and syntactic bootstrapping (Gleitman, 1990; Landau & Gleitman, 1985). We do not discuss these theories here because, in the case of SLA, it can be assumed that the first problem is not at issue: L2 learners come to the task with a highly developed (implicit) system of word classes in their L1. Nevertheless, the second problem—classifying newly learned words into the appropriate categories—applies to L2 learners as well.

There are several types of cues on which L1 and L2 learners can rely to categorize words into classes such as noun (N), verb (V), and adjective (Adj), among others. These cues are distributional and semantic.[2] Distributional cues include information about the absolute position of words in a sentence (e.g., first or last) as well as the relative position of words with respect to one another (i.e., lexical co-occurrence patterns). Crucially, distributional information also comprises grammatical morphemes: bound morphemes (inflections and derivations) and free morphemes (articles and auxiliary verbs). Gerken (2001) explained that "nearly every theory of language acquisition asserts that children use grammatical morphemes to assign syntactic category at some stage of development" (p. 155). For example, in English, the past tense suffix *–ed* attaches to verbs (e.g., *cooked*, *painted*) but not to nouns (**desked*, **fooded*) or prepositions (**fromed*, **ined*). Furthermore, words that take *–ed* to denote past tense generally accept *–s* to indicate present tense (e.g., *he cooked*, *he cooks*). In other words, there is morphological consistency across classes of items (Maratsos, 1988; Maratsos & Chalkley, 1980). The learner discovers that one of these characteristics predicts the other (and vice versa), and that possessing both characteristics is a good indicator that the word is a member of the category V. Additional cues are provided by free morphemes such as auxiliary verbs and articles: English verbs can be preceded by *do*-forms (e.g., *he doesn't cook*, *he did cook*), which distinguishes them from nouns that are often preceded by *the* or *a*.

Distributional information has proven to be a powerful mechanism for determining a word's category membership, as demonstrated by connectionist models (Finch & Chater, 1992, 1994), statistical networks of large language corpora (Redington & Chater, 1998), artificial language learning studies (Mintz, 2002), and studies of child-directed speech (Cartwright & Brent, 1997; Mintz, 2003; Mintz et al., 2002); for example, Mintz et al. demonstrated how the

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

acquisition of word class can proceed on the basis of the analysis of lexical cooccurrence (i.e., the links between words and their immediate context). Utilizing a child-directed language corpus, Mintz et al. found that even a very limited distributional context of one word before and after the target word was sufficient to cluster the majority of nouns and verbs into separate categories.

Finally, semantic cues can also contribute to the process of categorizing words into distinct classes, although the extent to which semantic information is used by children is controversial (see Gerken et al., 2005, for a review). The immediate problem with semantic information is that it is not a reliable predictor of word class because the correspondence between semantic and syntactic categories is not one-to-one. Consider, for example, the fact that *like* and *fond* are semantically similar, yet belong to different word classes. The only way to categorize them is to test their grammatical properties: *Like* behaves like a verb (it accepts *–ed* to denote past tense, and it can be preceded by forms of *do*), and *fond* behaves like other adjectives (it can be preceded by forms of *to be*). Even if semantic information is used at some point in the acquisition process to classify the prototypical members of each category (Bates & MacWhinney, 1982; Berman, 1988), it ultimately works in combination with more powerful distributional cues.

In the L2 context, using semantic information in assigning word class means finding the L1 equivalent. Positive transfer is expected to occur in the semantic domain because L1-L2 equivalents will belong to the same word class; for example, upon learning the word *mesa* "table" in Spanish, an English-speaking learner can simply rely on the L1 equivalent ("table") and correctly classify the new word as a noun. As Sunderman and Kroll (2006) explained, "two words that are translation equivalents, if they are otherwise unambiguous, are necessarily members of the same grammatical or lexical class" (p. 396).[3] However, a substantial number of words in English are ambiguous with respect to lexical class, which creates a problem for L2 learners.

Crosslinguistic Variation

The problem with making word class distinctions in one's L2 is that distributional regularities—both at the morphological and lexical levels—are language-specific. Knowing that English adverbs typically end in *–ly* does little to help a learner figure out which words are adverbs in Quechua, for example. Furthermore, languages differ in terms of the relative importance of morphology in assigning word class. Berman (1988) proposed that languages can be placed along a continuum: At one end are languages in which the surface form of words provides minimal cues to word class (e.g., Chinese), and at the other end are languages in which form and class are largely isomorphic (e.g., Hebrew, Latin). Here, we focus on English and Spanish and place these languages on the typological continuum proposed by Berman.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

According to Berman (1988), English is like Chinese in the sense that it depends primarily on syntactic properties of co-occurrence to assign word class. The most obvious example of this property is the presence of many identical verb and noun forms. For instance, the only way to determine if *drink*, *smile*, and *run* are nouns or verbs is by looking at the adjacent words in the sentence. This word-formation process is known as zero derivation. Nevertheless, Berman pointed out that English vocabulary of Latin origin displays considerable morphological cues to word class (e.g., *identify, identifiable, identification*). Another important word-formation process in Germanic languages such as English is compounding. This process yields frequent N-N combinations such as *dog trainer*, *property tax*, and *bike rack*, in addition to a seemingly infinite number of novel terms. Languages that do not have N-N compounding must resort to other mechanisms for expressing these terms. Thus, a *dog trainer* in Spanish is an *entrenador de perros* but never **perro entrenador*.[4] English has a wide range of compounding patterns in addition to the N-N type and most of them are very productive.

Spanish would be placed on the opposing end of the continuum described by Berman (1988) because of its rich system of inflections and derivations. Nouns, adjectives, and determiners are inflected for number and gender; verbs encode tense, mood, person, number, and aspect. In comparison with English, Spanish verbs have far more inflectional categories and affixes to mark them (e.g., there are 48 distinct simple inflectional forms according to Whitley, 2002). Furthermore, in terms of word formation, derivational morphology in Spanish is more productive than it is in English. Hence, whereas English relies on compounding, Spanish makes use of its abundant derivational affixes to create new words. The examples in (3) illustrate N-N compounding in English compared to Spanish adjectival suffixes.

(3) a. university professors—*profesores universitarios*
b. car accidents—*accidentes automovilísticos*
c. border patrol—*patrulla fronteriza*

Despite the rich derivational morphology of Spanish, some category overlap occurs between certain N-Adj pairs (e.g., *trabajador* "worker" or "hard-working," *político* "political" or "politician") due to the similarity of noun and adjective suffixes. Another kind of overlap has been noted between certain Adj-adverb (Adv) pairs such as *rápido* "quick" and *lento* "slow," but the same applies to English as well.

This brief linguistic sketch of Spanish and English is not meant to be exhaustive but rather serves to illustrate the relative importance of morphology as a cue to word class membership. In Spanish, a change in lexical category usually requires the appropriate change in derivational morphology, with the exception of some Adj-N pairs. English, on the other hand, makes extensive use of zero derivation resulting in identical N-V pairs (e.g., *drink* and *drink*) as well as V-Adj overlap (e.g., *open* and *open*). As a consequence,

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

for the English-speaking learner of Spanish, the problem will be not only to learn which Spanish grammatical morphemes mark particular classes of words but also to realize that the L1 equivalent is not always a reliable indicator of word class, given the pattern of zero derivation.

KEY FINDINGS FROM L2 STUDIES

Although little is known about the process of acquiring word classes in SLA, there are a number of studies that have addressed this problem indirectly. These studies can be divided into three types: (a) studies that have analyzed lexical errors in L2 learners' production (Whitley, 2004), (b) studies of L2 vocabulary that focus on learners' ability to produce appropriate derivatives (Morin, 2003, 2006; Schmitt & Zimmerman, 2002), and (c) psycholinguistic studies that tap into learners' underlying sensitivity to word classes (Sunderman & Kroll, 2006).

Whitley's (2004) analysis of lexical errors in the written production of upper division L2 Spanish learners highlights the significant problem of word class. Although the learners in his study were beginning to apply derivational morphology creatively (e.g., **sobrevivación* instead of *supervivencia* "survival"), many errors stemmed from zero derivation—that is, using "a single general form of a word in order to derive other forms without adding (or removing) derivational affixes" (p. 166). The most frequently occurring zero derivation errors involved verbs used as nouns, nouns used as adjectives, and adjectives used as nouns. To account for a wide range of word class errors that could not be explained by transfer, Whitley concluded that such errors reflect a "performance variable of student production from one moment to the next" (p. 171). Although Whitley presented a detailed description of word class production errors, the study leaves many questions unanswered about the nature of learners' difficulty with word classes in general.

Schmitt and Zimmerman (2002) examined English as a L2 learners' knowledge of word classes by testing their productive knowledge of derivatives that belong to the same word family. For example, if a learner demonstrates knowledge of the verb *indicate,* can we expect that learner to also produce the word *indication* when the context requires a nominal form? After a test of global vocabulary knowledge (Test of Academic Lexicon, adopted from Scarcella & Zimmerman, 1998), the participants were presented with four contextualized sentences for each prompt word, each of which required a different derivative form (e.g., *select, selection, selective, selectively*). The results indicated that participants had only partial knowledge of word families and were most often able to supply only two of the four target forms. Verb and noun forms were supplied more often, whereas adjective and adverb forms were generally less well known. Most interestingly, for the words rated as well known on the test of global vocabulary knowledge, only 65.5% of the corresponding derivative forms were produced. Schmitt and Zimmerman's study confirms

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

that learners have gaps in their knowledge of derivational morphology: Knowing one member of a word family does not guarantee knowledge of its related forms. The findings of Schmitt and Zimmerman are consistent with those of Schmitt and Meara (1997), who found that learners of English as a foreign language were able to produce only 15% of derivative suffixes in a written production task.

Recent research by Morin (2003, 2006) built upon the findings of Schmitt and Meara (1997) and Schmitt and Zimmerman (2002) by proposing a pedagogical intervention aimed at expanding L2 learners' vocabulary through the explicit teaching of Spanish derivational morphology. The participants in Morin's studies completed tests of receptive and productive knowledge of Spanish word families. We focus on the findings of the receptive task in which learners completed sentences with one of four forms belonging to the same word family (e.g., the noun *estudiante* "student," the noun *estudio* "study," the adjective *estudiantil* "student," the adverb *estudiosamente* "studiously"). Morin (2003) found that first- and second-semester learners of Spanish had very limited receptive morphological knowledge; the mean scores of each group were 3.4 and 7.3 (out of a possible 20 points) prior to treatment. Although both groups made gains after the instructional treatment (the posttest mean scores were 12), these results suggest that beginning-level learners are sometimes unable to fill a syntactic slot with the appropriate derivative, even when presented with the correct option in a multiple-choice format. Morin (2006) repeated the study with third-semester learners but added a metalanguage component to the receptive task. Thus, in addition to choosing the form to complete each sentence, learners had to indicate whether the chosen form was a noun, a verb, an adjective, or an adverb. The results indicated that prior to treatment, the learners scored slightly better than chance (the mean was 21.7 of 40). However, because the scoring of the task conflates tacit knowledge of word classes (choosing the correct form) and metalinguistic knowledge (indicating the corresponding label), it is impossible to know where learners were losing points without further analysis of the data.[5] Furthermore, in both studies, the receptive task was limited to five verbs and their related derivational forms and, therefore, does not permit a detailed analysis of which items were the most problematic for the learners.

A recent study of lexical activation among learners of L2 Spanish by Sunderman and Kroll (2006) investigated learners' implicit sensitivity to word classes. In a complex research design, Sunderman and Kroll presented learners with Spanish-English translation equivalents (e.g., *cara* and "face") as well as different types of critical distracters in order to measure the degree of interference from form-related and meaning-related words. An interesting design feature was the inclusion of distracters that were either a match or a mismatch with the target word with respect to word class; for example, the noun *cara* was paired with the noun *fact* in the matched condition, whereas in the mismatched condition, it was paired with the adjective *fast*. If learners

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

are sensitive to the grammatical class of *cara*, they should be able to reject *fast* more rapidly than *fact*. The results suggest that both less and more proficient learners were able to use word class as a cue to reject words that had a high degree of form-related similarity (e.g., *cara* and *fast*). Sunderman and Kroll concluded that word class information in the L2 might be available to even less proficient learners, in spite of the fact that the task presented words in isolation (i.e., there were no distributional cues to word class). However, their data also show that meaning-related distracters (e.g., *cara* and *pretty*) resulted in interference regardless of matched or mismatched word class. Although Sunderman and Kroll did not test items from the same word family (e.g., *feliz* and *happiness*), their results suggest that such pairs might also result in interference despite the word class mismatch.

To summarize, word classes in SLA have been examined only indirectly—that is, from the perspective of acquiring new words and applying derivational affixes. The study of derivatives is warranted because derivational morphemes can change the lexical category of a base word (e.g., adding the suffix *–tion* to the verb *indicate* creates the noun *indication*). However, it is important to remember that inflectional morphemes also play a crucial role in identifying a word as a member of a particular category. Although SLA research has a tradition of separating the study of inflectional and derivational morphology, the problem of word class requires an integrated approach.[6] In addition to morphology, the role of syntactic cues (i.e., lexical co-occurrence properties) must not be overlooked. As such, tasks must be designed with sentential contexts so that learners can use distributional information as well as morphological cues.

RESEARCH QUESTIONS

The present study expands on the initial findings of Whitley (2004) and Morin (2003, 2006) to determine the nature of L2 learners' difficulties with word classes. It is part of a larger study on the acquisition of word classes in which productive, receptive, and introspective data were collected. Here, we present the development and implementation of a receptive task that targeted specific word class contrasts (e.g., N-V, Adj-V) in a variety of syntactic contexts. The development of a reliable instrument to answer the two research questions was an important goal of the study. The research questions that guided the design of the study are as follows:

1. Are word class confusions limited to production, or do L2 learners also struggle to distinguish among word classes in the receptive mode?
2. Are certain kinds of contrasts (e.g., N-V) particularly difficult for learners to perceive? If so, what is the effect of surface morphology on learners' performance?

With respect to the first research question, we hypothesize that word class confusions will show across-the-board effects: Learners will make errors in

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

both productive and receptive tasks. If this hypothesis is confirmed, then it would indicate that word class errors are not merely a performance phenomenon as Whitley suggested. For the second research question, we predict that N-Adj forms will be the most difficult to distinguish: The learning task is more complex for L2 learners because there is some category overlap in Spanish. Furthermore, we predict that learners will demonstrate greatest success with words whose derivational suffixes clearly mark word class (e.g., *–ción* as a marker of nouns).

To test our hypotheses, we sought to develop a receptive task that met two basic requirements: First, it should be representative of the kinds of word class errors actually produced by learners, and, second, it must give learners options for choosing among words that belong to different lexical classes without invoking labels such as noun, verb, and adjective, among others.

METHOD

Development of the Instrument

A preliminary step in developing a reliable instrument was to gather word class errors produced by L2 learners. Given that Whitley (2004) documented many such errors in an in-class expository essay, we followed a similar procedure and asked learners of Spanish from a third-semester class (n = 74) to write responses to open-ended topics chosen in consultation with the course instructors. As an example, one of the writing prompts asked learners to elaborate on the concept of good manners and to give examples from their experience. Learners were given 15 min to write on the assigned topics during three different class periods. They were encouraged to write as much as possible in the allotted time and to expand on the given topic in any way they wanted. They had no recourse to dictionaries or any other materials while performing the written production tasks. Given the open-ended nature of the written tasks, learners could certainly avoid what they perceived to be difficult grammatical constructions or unfamiliar lexical items. Nevertheless, we deemed three written responses from each learner to be sufficient for gathering a varied corpus of word class errors.

In coding the written data, we operationalized a word class error as the use of a target language word in an inappropriate grammatical frame. These correspond to the category of zero-derivation errors in James (1998) and Whitley (2004)—that is, the use of a single form of a word regardless of syntactic context. The list of criteria for coding the written data (see Appendix A) was used by both researchers who exhibited a 94% agreement rate in error identification. Of the total errors identified, there was 100% interrater agreement on the type of error (noun used as an adjective, verb used as a noun, etc.).

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

These written production tasks yielded a total of 132 word class violations. These were further categorized in terms of type of error; for example, a verb form such as *abra* "open" (V, present, subjunctive) instead of the corresponding adjective *abierta* was classified as a verb used as an adjective. Note that verb refers exclusively to finite verbs forms.[7] Representative examples from the learner-generated corpus are provided in Appendix B with the correct form provided in brackets. The written production tasks revealed that certain error types were much more prevalent than others. Finite verb forms constitute the main source of error, accounting for 59% of the total word class violations. Particularly common were verbs used as nouns, as infinitives (Inf), and as adjectives. Adjectives were also frequently used as nouns. Interestingly, the majority of word class errors could not be traced back to the L1, as illustrated in (4).

(4) *Es más bueno por el económico de los países.*
"It's better for the economic of the countries."

Note that the translation equivalent of the Spanish word *económico* results in a word class violation in English as well.

After coding the written production data, we designed a receptive task based on the four most common error types: a verb used as an adjective (condition 1), a verb used as a noun (condition 2), a verb used as an infinitive (condition 3), and an adjective used as a noun (condition 4). This receptive task is best described as a forced-choice task (FCT) that consisted of 32 target items and 32 fillers. Each target item consisted of a sentence taken directly from learners' written production.[8] An example of a test item is provided in (5).

(5) *Mis padres no me permitieron comer con la boca (abra / abierta).* neither
"My parents didn't let me eat with my mouth (open / open)."

The item in (5) was taken from one learner's written essay about good manners in which he or she used the verb form *abra* "open" in this slot instead of the required adjective form *abierta* "open." In this way, the FCT reproduces the original word class error but provides the correct form alongside it. The use of real data in the elaboration of the FCT is important because it allows us to work with lexical items and constructions familiar to learners rather than contrived sentences generated for research purposes. We also checked the first- and second-year textbooks used in this university's curriculum in order to limit the FCT to words likely to be used and presented in classroom contexts.[9] All of the target words are found in either one or both texts and, moreover, the majority are presented as active vocabulary.

The items on the FCT target contexts that were the source of word class errors in the learners' written production. Another option in designing the FCT would have been to include some contexts in which learners demonstrated adequate knowledge of word class distinctions, thereby expanding the scope of the task.[10] We decided to target only the problematic contexts

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

in order to test our hypothesis that word class violations would show across-the-board effects. Finally, the FCT was designed to include both inflected and derived forms (e.g., *sientes* "you feel" vs. *sentimientos* "feelings") given the fact that both inflections and derivations serve as cues to word class.

The instructions for the FCT asked learners to circle the correct form or forms that could complete the sentence. If the learner felt that neither form was appropriate, the learner could check *neither* and write in an alternative. This format gave learners four options: choosing either one of the forms, both forms, or neither form. The advantage of this format over a simple binary one is that it should reduce the likelihood of guessing. The order of the words in parentheses (e.g., *abra / abierta* "open") was randomized so that the correct form was not always in the same position. The target items included in the FCT are provided in Appendix C, in which they are grouped by condition (eight items in each condition). On the version administered to the learners, these items were randomized with the fillers, which mirrored the options for the target items. It was particularly important to include fillers in which both options were possible but with forms that did not radically differ from one another. For such items, we presented nouns with invariable adjectives (e.g., *mi hermano / hermana mayor* "my older brother / sister") and verbs with a tense and aspect contrast (e.g., *me ayuda / ayudaba* "he helps / helped me"). The task was administered in paper-and-pencil format and was not timed.

To ensure the reliability of the instrument, a measure of internal consistency was computed. Such an analysis enables one to determine the consistency of responses on the target items, which reflects how well the items measure the same construct (in this case, receptive knowledge of word classes). The 32 target items yielded an alpha-reliability coefficient of .82, indicating a good level of interrelatedness among the items.

Participants

A total of 240 English-speaking learners of Spanish at a large public American university participated. Approximately half (n = 122) were enrolled in a third-semester course, which they were taking to satisfy a general requirement. Seventy-four of these learners had participated in the development of the instrument via their writing samples gathered earlier in the semester. Another group (n = 118) consisted primarily of Spanish majors and minors who were enrolled in a third-year advanced grammar course. Students were placed in the appropriate level based on the results of a university placement test that includes reading comprehension, grammar and vocabulary knowledge, and writing ability. Both groups of learners had studied Spanish in the classroom context; that is, they had not participated in the university's study-abroad programs at the time of data collection.[11]

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Both groups were tested during the last week of the semester, after having completed 15 weeks of Spanish language instruction in their respective courses. Both courses are communicatively oriented, but the third-year course stresses accuracy in areas of Spanish grammar typically difficult for non-native speakers (e.g., prepositions, pronouns, subordination). As far as explicit instruction on word classes, this is not a focus of instruction at any point in the curriculum, and thus we assume that it did not take place in either course.[12] Of course, learners are exposed to the labels associated with word classes (e.g., N, V, Adj), but this does not help in figuring out their distribution. Even if some distributional information is explicitly presented, as in the case of adjectives (i.e., students are taught that adjectives usually follow nouns), it still leaves the problem of figuring out which words in the target language are nouns and which are adjectives. In this sense, learners are left to their own devices: They must abstract the regularities from the L2 input or rely on transfer from the L1.

RESULTS

The Forced-Choice Task

On each of the target items (k = 32), participants received 1 point for choosing the correct form and 0 for any of the remaining possible responses (i.e., choosing the wrong form, choosing both forms, or marking *neither*). The first set of statistical analyses was carried out on each participant's total score (with a maximum of 32) on the FCT. These results are shown in Table 1. An independent samples t test confirms that the difference between the two groups is statistically significant, $t(238) = -7.35$, $p < .001$.

The next step was to consider the results for each of the four conditions. Recall that there were 8 items in each condition, for a total of 32 target items. Mean scores for each condition for both groups are listed in Table 2. To determine if there are significant differences in learners' performance on these four conditions, a two-way repeated measures ANOVA was conducted. The within-subjects factor was condition, and the between-subjects factor was group. There was a significant main effect for condition, $F(3, 714) = 65.9$, $p < .001$, and a significant main effect for group, $F(1, 238) = 54.08$, $p < .001$, but no interaction between group and condition. Post hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons. Results indicate that performance on condition 4 was significantly lower than performance on the other three conditions ($p < .005$). Significant differences were also found between condition 1 and the other conditions ($p < .005$). The comparison between conditions 2 and 3 did not reach statistical significance.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Table 1. Descriptive statistics for total score on FCT

Learner	*N*	*M*	*SD*
Third-semester	122	20.88	5.25
Third-year	118	25.55	4.55

Table 2. Mean scores per condition and learner group

	Third-semester		Third-year	
Condition	*M*	*SD*	*M*	*SD*
1	5.88	1.62	7.03	1.38
2	5.29	1.65	6.25	1.49
3	5.51	1.93	6.64	1.57
4	4.20	1.76	5.63	1.67

Item Analysis

To determine which particular items within each condition were most difficult for the learners, we performed an item analysis. Table 3 shows the means, standard deviations, and part-whole correlations for each target item. The means presented in Table 3 indicate that there were seven target items (containing the word pairs *intenta–intención* "he or she intends–intention," *tienen–tener* "they have–to have," *pasar–pasan* "to spend–they spend," *corte–corto* "cut–short," *fuerte–fuerza* "strong–strength," *cómodo–comodidad* "comfortable–comfort," and *cariñoso–cariño* "affectionate–affection") with mean scores around .5 or less, which means that learners' performance on these items was at chance level. (For the complete sentences in which these word pairs appeared, see Appendix C.) The last four of these pairs involve an adjective and noun that belong to the same word family, illustrating the kind of contrast that yielded significantly lower mean scores in condition 4. In contrast to these pairs, several items yielded very high mean scores (.9 or above). These were items that contained the word pairs *necesario–necesito* "necessary–I need," *divierte–divertido* "amuses–fun," *tiene–tener* "he or she has–to have," and *comer–come* "to eat-he or she eats." For these items, the large majority of learners had no difficulty choosing the correct form.

Given our fairly large cross-section of learners, we were interested in finding out which items were most frequently missed by the learners who registered high scores on the task overall. To this end, we ran another item analysis only with the learners who scored a total of 27 or higher on the task (i.e., the most accurate performers in the sample). The 76 learners in this subsample had

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Table 3. Item analysis

Word pairs	*M*	*SD*
Abra–abierta "open (subjunctive)–open"	.78	.41
Necesario–necesito "necessary–(I) need"	.90	.29
Cierra–cerrada "(he/she) closes–closed"	.78	.42
Válidas–vale "valid–worth"	.71	.45
Importa–importante "matter–important"	.72	.45
Activos–activan "active–(they) activate"	.73	.44
Importantes–importan "important–matter "	.88	.33
Divierte–divertido "amuse–fun"	.94	.24
Mentiras–mientes "lies–(you) lie"	.75	.43
Intenta–intención "(he/she) intends–intention"	.55	.50
Disciplina–disciplino "discipline–(I) discipline"	.59	.49
Sientes–sentimientos "(you) feel–feelings"	.77	.42
Pelean–peleas "(they) fight–fights"	.71	.45
Mientes–mentiras "(you) lie–lies"	.80	.40
Trabajo–trabaja "job–(he/she) works"	.74	.44
Respuesta–responde "answer–(he/she) answers"	.85	.36
Miente–mentir "(he/she) lies–to lie"	.73	.45
Tiene–tener "(he/she has)–to have"	.90	.31
Comer–come "to eat–(he/she) eats"	.91	.29
Abrir–abre "to open–(he/she) opens"	.86	.35
Miente–mentir "(he/she) lies–to lie"	.85	.36
Decir–dice "to say–(he/she) says"	.79	.41
Tienen–tener "(they) have–to have"	.52	.50
Pasar–pasan "to spend–(they) spend"	.51	.50
Divorcio–divorciado "divorce–divorced"	.84	.37
Corte–corto "cut–short"	.35	.48
Fuerte–fuerza "strong–strength"	.53	.74
Emocionados–emociones "excited–emotions"	.76	.43
Comodidad–cómodo "comfort–comfortable"	.48	.50
Cariñosa–cariño "affectionate–affection"	.57	.49
Económico–economía "economical–economy"	.75	.43
Felicidad–feliz "happiness–happy"	.65	.48

the lowest mean scores on the following three items: *intenta–intención* "he or she intends–intention" (*m* = .67), *corte–corto* "cut-short" (*m* = .64), and *comedidad*–comodo "*comfort–comfortable*" (*m* = .66).

Some patterns in the data reveal inconsistent responses: Learners were highly accurate with the V-Inf contrast with the word pairs *tiene–tener* "he or she has–to have" and *comer–come* "to eat–he or she eats" (mean scores of .9 or higher), whereas they struggled with the very same contrast with the word pairs *tienen–tener* "they have-to have" and *pasar–pasan* "to spend–they spend" (mean scores of .52 and .51, respectively). Because the word pairs *tiene–tener* "he or she has–to have" and *tienen–tener* "they have–to have"

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

both contrast the infinitive *tener* with a finite form, these items, reproduced in (6) and (7), make for an interesting comparison.

(6) *Un ejemplo de buenos modales es (tiene / tener) respeto por otras personas.*
"An example of good manners is (have / to have) respect for other people."

(7) *Yo creo que es importante para los hijos (tienen / tener) contacto con sus parientes.*
"I believe that it is important for children (have / to have) contact with their relatives."

Learners performed significantly better on the word pair *tiene–tener* "he or she has–to have" (m = .9) than on the word pair *tienen–tener* "they have–to have" (m = .52), as indicated by a paired-samples *t* test, $t(239) = 11$, $p < .001$. These results suggest that the intervening plural noun phrase los *hijos* "the children" in the item containing the word pair *tienen–tener* "they have–to have" led many learners to choose the finite form rather than the infinitive. We will now consider these results in light of the various morphological and syntactic cues that affect learners' ability to discriminate among word classes.

DISCUSSION

Main Findings

With respect to the first research question—whether word class confusions are limited to production—the results indicate that learners also experience some problems with word classes in the receptive mode. The results from the FCT confirm our first hypothesis. The mean score on this task was just above 20 (out of 32) for the third-semester group, which indicates that these learners are able to choose the required word class approximately two thirds of the time. As expected, the third-year learners performed significantly better (M = 25.55) but still have difficulty with some items on the task. These results are important because they challenge the view that errors in word class are simply a performance problem, as proposed by Whitley (2004). In contrast, our results indicate that learners are often unable to recognize the appropriate form that fits the grammatical slot, even when it is presented to them in multiple-choice format.

Nevertheless, these difficulties with word class in the receptive mode are not generalized to all contexts and to all items. In response to our second research question, we observe that, given the four conditions, learners were least accurate in distinguishing between adjective and noun forms and most accurate with the V-Adj contrast. The finding that L2 learners have more difficulty distinguishing between adjective and noun forms may be due to properties

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

of the input. First, some adjectives and nouns in Spanish are identical in form (i.e., they are zero derived). Second, there is some positional overlap between adjectives and nouns due to the noun-drop phenomenon (e.g., *prefiero el rojo* "I prefer the red one"), which is highly productive in Spanish (White, Valenzuela, Kozlowska-Macgregor, & Leung, 2004). Finally, the position of adjectives vis-à-vis nouns is not fixed, which means that positional cues are not always sufficient for distinguishing between the two classes of words.

The item analysis reveals that Adj-N pairs that were most difficult for learners to distinguish were *fuerte–fuerza* "strong–strength," *corte–corto* "cut–short," *cómodo–comodidad* "comfortable–comfort," and *cariñoso–cariño* "affectionate–affection." Some aspects of these findings can be easily explained by ambiguous surface morphology. Consider the N-Adj pair *corte–corto* "cut–short" that differs only in the final vowel (*e* or *o*). It is not surprising that learners confuse these forms, as both *–e* and *–o* are ambiguous markers that appear across word classes. Similarly, the ambiguity of *–a* and *–o* endings in V-N pairs makes it challenging to distinguish between forms like the noun *trabajo* "job" and the verb *trabaja* "he or she works" (m = .74) or the noun *disciplina* "discipline" and the verb *disciplino* "I discipline" (m = .59). This sort of word class confusion is bound to occur in L2 Spanish because the *–a* and *–o* endings that distinguish word classes are also markers of grammatical gender, and L2 research has repeatedly illustrated learners' variable performance with gender marking (see, e.g., Alarcón, 2006; Bruhn de Garavito & White, 2002). Thus, we suggest that word class confusions between pairs such as the noun *trabajo* "job" and the verb *trabaja* "he or she works" be viewed as another manifestation of gender (mis)assignment.

The apparent confusion between the adjective *cómodo* "comfortable" and the noun *comodidad* "comfort" was unanticipated, however, because the *–dad* suffix is an unmistakable marker of nouns in Spanish. The same is true for the adjective *cariñosa* "affectionate" and the noun *cariño* "affection," in which the first form is unambiguously marked with the adjective suffix *–oso*. Recall also that the V-N pair *intenta–intención* "he or she intends–intention" yielded low mean scores, even among learners who otherwise performed well on the task. Confusions between forms that are clearly morphologically marked suggest that learners do not recognize these derivational suffixes as markers of word class. In other words, they are lacking syntactic morphological knowledge (Tyler & Nagy, 1989) that involves knowing that a particular derivational suffix marks a word for a specific lexical category. Although *–ción* is a clear marker of nouns in Spanish and shares a close resemblance with the English suffix *–tion*, many learners incorrectly chose the verb *intenta* "he or she intends" presumably because it ends in an *–a* and therefore could potentially agree with the preceding word in the sentence (*mala* "bad").[13] There could be an instruction effect here: Learners in classroom contexts are repeatedly taught, from the first weeks of study, that feminine nouns end in *–a* (e.g., *casa* "house") and masculine nouns end in *–o* (e.g., *teléfono* phone"). Although not an inaccurate generalization, it neglects the crucial fact that

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

other classes of words also end in *–a* and *–o*. As Harris (1991) pointed out, the suffixes *–o* and *–a* are word markers rather than gender markers because they are not limited to lexical items that have gender. However, classroom learners are encouraged to pay attention to *–a* and *–o* endings as gender markers at the expense of morphological cues to word classes (e.g., *–ción* or *–dad* as markers of nouns).

If learners lack crucial knowledge of the relationship between suffixes and word class, they should be able to compensate for this by relying on properties of lexical co-occurrence. Recall that theoretical accounts of word class acquisition give great importance to distributional or positional cues. Consider the learners who selected the adjective *feliz* "happy" in the item *para tener feliz / felicidad* "in order to be happy / happiness." The mean accuracy on this item was .65, which means that 84 of the 240 participants chose the adjective to fill a slot that can only be occupied by a noun. It could be argued that these learners think *feliz* "happy" is a noun; that is, they do not recognize *feliz* "happy" as being exclusively an adjectival form. However, the syntactic contexts in which *feliz* "happy" appears should prevent this from happening; for example, classroom learners of Spanish will regularly hear this word in fixed expressions such as *feliz navidad* "merry Christmas" and *feliz cumpleaños* "happy birthday," in addition to sentences in which *feliz* "happy" is preceded by a copular verb (*ser* or *estar* "to be"). These are clear distributional signals that the form is an adjective. Thus, it is more likely that these learners did not recognize the grammatical frame *para tener* "in order to have" in the item *para tener feliz / felicidad* "in order to have happy / happiness" as requiring a noun. A plausible reason is the erroneous syntactic analysis of common expressions with *tener* followed by a noun, which are translated into English as *be* and an adjective (e.g., *tener calor* "to be hot"). If the learner had analyzed such expressions as *tener* followed by an adjective, then **tener feliz* is a perfectly logical, albeit incorrect, construction. In other words, there is reason to view this as a syntactic error, coupled with the inability to recognize *felicidad* "happiness" as the noun form that belongs in that slot.

Further evidence for the syntactic basis of some word class errors can be found in learners' inconsistent performance with the V-Inf contrast. We assume that all the learners were able to recognize Spanish infinitives by their morphological marking (the suffix *–r* after the theme vowel). This is a reasonable assumption because the infinitive is the citation form of the verb (i.e., the form that appears in textbook vocabulary lists and in dictionaries). Therefore, the confusion between finite verbs and infinitives is not a morphological problem but rather one of not knowing the distinct distribution of both classes of words. Unlike finite verbs, infinitives in Spanish are nounlike and, as a result, can appear in canonical noun positions (e.g., as subject, object, or object of a preposition). If learners are not aware of these distributional regularities, they may overuse the finite form in an attempt to establish agreement.

The discussion of the results thus far suggests that word class violations stem from at least two sources: learners' incomplete knowledge (a) of L2

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

derivational suffixes and their relationship to word classes and (b) of L2 distributional regularities. In this sense, we maintain that word class errors are not merely a performance phenomenon but are reflective of a competence deficit. Crucially, our use of the term "competence" does not refer to abstract grammatical representations, such as those that have been postulated by Universal Grammar–oriented researchers. Word categorization is a quintessential learning problem, as Ellis (2003) described:

> Whether there are linguistic universals or not, there is still a logical problem of syntactic acquisition. Identifying the syntactic category of words must primarily be a matter of learning because the phonological strings associated with words of a language are clearly not universal. Once some identifications have been made, it may be possible to use prior grammatical knowledge to facilitate further identifications. But the acquisition of relevant phrase structure grammar requires knowledge of syntactic word class in the first place (pp. 78–79).

Even if a competence problem underlies L2 learners' difficulties with word class, this does not mean that the deficit is permanent or nonintegratable in Jiang's (2007) words. On the contrary, some learners in the sample registered nearly perfect scores on the FCT (e.g., 31 learners scored 30 or higher). This indicates that knowledge of word classes develops with increased exposure to the L2 and is eventually fully integrated. The fact that many learners are able to do this successfully after 3 years of study, despite the limited quantity and quality of classroom input, suggests that implicit learning of word classes can (and does) take place in instructed SLA. Our results are in line with those of Morin (2006), who found no effect of explicit instruction on receptive knowledge of word classes but suggested that such knowledge develops incidentally.

Frequency Effects and Pedagogical Applications

An additional factor that should be considered when interpreting the data from the FCT is word frequency. It could be that a learner, when faced with two related forms such as the adjective *feliz* "happy" and the noun *felicidad* "happiness," simply chooses the one that is more frequent in the input. To ascertain the possible effects of word frequency, we conducted a post hoc analysis of the target forms that appeared on the FCT by consulting a recent frequency dictionary of Spanish (Davies, 2006) that provides a rank-ordered list of the 5,000 most frequent words (based on a 20-million-word corpus). Because the frequency dictionary lists lemmas rather than exact word forms, it is impossible to compare finite verb forms such as *miente* "he or she lies" or *mientes* "you lie" and the corresponding infinitive *mentir* "to lie." Nevertheless, it allows for a precise comparison of some of the most problematic items on

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

the FCT. In the case of the adjective *feliz* "happy" and the noun *felicidad* "happiness," the adjective is indeed more frequent than the noun (ranked at 956 and 1,892, respectively). However, in other cases, frequency favors that word that was chosen less often by the L2 learners. For example, the noun *cariño* "affection" is more frequent than the adjective *cariñoso* "affectionate" (ranked 2,246 and 4,033, respectively); similarly, *fuerza* "strength" is more frequent than *fuerte* "strong" (ranked 255 and 1010 respectively). Finally, other word pairs were almost identical in terms of their rank frequency (e.g., the noun *trabajo* "work" ranked 145, and the verb *trabajar* "to work" ranked 183). In sum, the relationship between word frequency and learners' performance on the FCT is inconsistent. Although the possibility of frequency effects in our data cannot be ruled out, there is evidence to suggest that other factors (i.e., incomplete syntactic and morphological knowledge) are responsible for word class difficulties. Crucially, the frequency of a given form does not explain why learners opt for a syntactic frame that requires a different word class; for example, during production, it could be that a L2 learner uses *feliz* "happy" because it is the form that the learner has heard more frequently. However, frequency does not explain why the learner opts for the syntactic frame *in order to have* followed by Adj. This frame, taken directly from our written production data, could have been constructed with the verb *ser* "to be," thus allowing the learner to use *feliz* "happy" appropriately as an adjective. This highlights the syntactic nature of many word class errors, despite the possibility of frequency effects that may predispose learners to using particular forms.

To conclude, we provide some suggestions regarding pedagogical interventions that may improve learners' ability to distinguish among word classes. The notion of a pedagogical treatment for word categorization is somewhat counterintuitive if one assumes that categorizing words proceeds on the basis of distributional analysis (and is thus largely implicit). Nevertheless, the findings of the current study point to several possibilities that might improve L2 learners' strategies for coping with word class contrasts. One option is to provide instruction on derivational suffixes and how these relate to word class (see Morin, 2003, 2006). Such instruction should aim not only at expanding knowledge of word families (e.g., *fuerte* "strong," *fuerza* "strength," *fortalecer* "strengthen") but also highlight the reliable morphological markers of certain word classes (e.g., words ending in *–dad* and *–ción* are nouns). More importantly, this type of instruction should be coupled with a focus on distributional regularities; for example, it may be useful to teach students that only nouns can follow prepositions or that infinitives and nouns are similar in terms of distribution.

A related matter is the order of presentation of active vocabulary in textbooks. In our examination of first- and second-year texts, we found that these textbooks made an attempt to group words into families, but their presentation was largely inconsistent. As an example, the intermediate-level text (Bretz, Dvorak, Kirschner, Bransdorfer, & Kihyet, 2004) presents the noun *la disciplina*

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

"discipline" and the verb *disciplinar* "to discipline" together but fails to present the nominal counterpart to the adjective *cariñoso* "affectionate." The first-year text (Terrell, Andrade, Egasse, & Muños, 2005) groups active vocabulary thematically (e.g., places, holidays) and by word class, but words belonging to the same family are often separated by many chapters; for example, the adjective *feliz* "happy" appears in the first chapter but the noun *felicidad* "happiness" is not presented until chapter 15. This confirms Whitley's (2004) observation that classroom learners have acquired

> whole hosts of loose lexical items such as *triste* "sad" and *aumentar* "increase," as commonly presented in textbook vocabulary lists, but not the connections to *tristeza* "sadness" or *el aumento* "the increase" or the adequate means of generating these connections on their own (p. 169).

Thus, it is up to the instructor to explicitly draw learners' attention to word families and the distinctions among related nouns, verbs, and adjectives.

CONCLUSION

The process for categorizing newly learned words into classes requires parsing the input of a variety of morphological, syntactic, and semantic cues. Often, the automaticity of this process is taken for granted, and the assumption is made that learners, upon encountering new words in the input, will categorize them correctly as nouns, verbs, and adjectives, among others. The findings presented here indicate that this is not always the case and that L2 learners have difficulty distinguishing among semantically related forms. The data also reveal that the N-Adj contrast is most difficult for L2 learners of Spanish and that unambiguous derivational suffixes (e.g., *–ción*, *–dad*) are not necessarily recognized by learners as markers of word class. We interpret learners' difficulties with word class as a competence deficit stemming from limited syntactic and morphological knowledge.

An important objective of the study was to develop an instrument that could be expanded or refined in future research. Testing word classes presents a series of challenges because of the confounding variable of metalanguage. Previous studies (Morin, 2006; Schmitt & Zimmerman, 2002) have used grammatical terminology as part of the testing instruments, which renders the findings difficult to interpret as either reflecting implicit knowledge of word classes or the ability to attach labels to them. Psycholinguistic tasks, such as the ones used in Sunderman and Kroll (2006), are perhaps the best way of tapping implicit sensitivity to word classes without involving metalanguage. Specifically with regard to lexical activation experiments, it would be interesting to include distracters from the same word family to measure the degree of interference. Another promising direction for future research would be to determine if a learner's ability to make word class distinctions correlates

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

with other components of L2 proficiency, specifically grammatical ability. Indeed, one limitation of the present study is that the findings are limited to group data, and, therefore, we cannot make any claims about individual differences.

In conclusion, this study provides insight into the nature of L2 learners' difficulties with word classes, a topic that had only been indirectly addressed in the L2 literature, primarily in studies of vocabulary acquisition. However, if learners struggle to discriminate among word classes in a receptive task due to incomplete knowledge of L2 distributional patterns and morphological marking, the problem of word class is not strictly one of vocabulary learning. Word classes are, in the end, the building blocks of syntax (hence the term *syntactic categories*). In this way, understanding the mechanisms of word classification in SLA is a fundamental part of describing learners' grammatical development in general. Further research is needed to shed light on which cues L2 learners rely on—or ignore—as they categorize newly learned words into distinct classes.

(*Received 23 January 2008*)

NOTES

1. We adopt Jiang's (2004, 2007) definition of competence as integrated knowledge, which is operationalized as "the information that a learner can retrieve and put to use without paying attention to it" (2007, p. 2). For word classes, this involves knowing which types of words co-occur with others and which grammatical morphemes can appear with certain classes of words.

2. For L1 acquisition, phonology and prosody can also provide important cues (Durieux & Gillis, 2001). As an example, English nouns are more likely than verbs to carry stress on the initial syllable (see Kelly, 1996, for additional evidence of the correlation between phonology and grammatical class). Recent accounts of word class acquisition include phonological properties under the rubric of distributional information.

3. We stress that these must be exact equivalents for the L2 learner to benefit from positive transfer. There is a great deal of crosslinguistic variation that we cannot fully address in the scope of this article (see Baker, 2003, for a detailed account). A particular semantic concept may be assigned to different word classes depending on the language; for example, some languages (e.g., Polish, Hebrew) have a verb *to be late*, whereas English has only the adjective. In Quechua, words that are adjectives in English (e.g., *red* and *tall*) have the morphosyntactic properties of nouns (Dixon, 1982). In sum, semantic information cannot always be used to determine the word class of an item in the L2.

4. There are some N-N compounds in Spanish (e.g., *coche bomba* "car bomb"), but this is not a productive pattern, as in English (cf. Piera, 1995). For SLA studies on N-N compounding, see Slabakova (2002) and Liceras and Diaz (2000).

5. R. Morin (personal communication, June 14, 2006) suggested that the stronger learners were usually able to choose the correct form but lost points by misidentifying the part of speech. The weaker learners, on the other hand, left many items blank (the instructions said not to guess) but also had difficulty in identifying the form as a noun, verb, adjective, or adverb.

6. Although inflectional morphology has been extensively examined in studies of L2 grammar, derivational morphology has mainly received attention from vocabulary researchers (e.g., Schmitt & Zimmerman, 2002).

7. This classification reflects the fact that only finite verbs are inflected for person, number, tense, aspect, and mood. Infinitive forms, on the other hand, function primarily as nouns in Spanish grammar and have a distinct distribution.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

8. Some minor changes were made to the learners' sentences in order to fix extraneous grammatical errors such as tense, gender, and preposition use (e.g., *por* and *para* "for"). Orthographic errors were also corrected. However, every effort was made to preserve the original sentence structure and lexical choices made by the learner.

9. The first-year text was *Dos Mundos* (Terrell et al., 2005), and the second-year text was *Avance* (Bretz et al., 2004).

10. This alternative was suggested by an anonymous *SSLA* reviewer. We recognize that the FCT, as we have designed it, is only one method for investigating the problem of word class. Additional methodological options are discussed in the conclusion.

11. The third-year grammar course is a prerequisite for this university's study-abroad programs in Spanish. Therefore, these learners were tested prior to any study-abroad experience. The only exceptions to this profile are two female students who had spent several months in Mexico and three heritage speakers, but their data were excluded from the analysis. As expected, the heritage speakers registered a perfect score on the task.

12. One of the researchers was the instructor of the third-year grammar course, and the other has taught the third-semester course numerous times. Given our experience with these courses, we can assert that word classes are not a focus of explicit instruction.

13. Our introspective data, which we do not discuss here for reasons of space, confirms that learners are paying attention to *–a* and *–o* endings as markers of gender agreement; for example, many learners mistakenly perceived *cariñosa* "affectionate" to be the feminine form of the noun *cariño* "affection."

REFERENCES

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

Alderson, J. C., Clapham, C. M., & Steel, D. (1997). Metalinguistic knowledge, language aptitude, and language proficiency. *Language Teaching Research*, *1*, 93–121.

Baker, M. C. (2003). *Lexical categories: Verbs, nouns, and adjectives*. New York: Cambridge University Press.

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, *11*, 17–34.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & R. L. Gleitman (Eds.), *Language acquisition: The state of the state of the art* (pp. 173–218). New York: Cambridge University Press.

Berman, R. A. (1988). Word class distinctions in developing grammars. In Y. Levy, I. M. Schlesinger, & M. Braine (Eds.), *Categories and processes in language acquisition* (pp. 45–72). Mahwah, NJ: Erlbaum.

Braine, M. (1987). What is learned in acquiring word classes: A step toward an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 65–87). Mahwah, NJ: Erlbaum.

Bretz, M. L., Dvorak, T., Kirschner, C., Bransdorfer, R., & Kihyet, C. (2004). *¡Avance!* New York: McGraw Hill.

Bruhn de Garavito, J., & White, L. (2002). The second language acquisition of Spanish DPs: The status of grammatical features. In A. T. Pérez-Leroux & J. M. Liceras (Eds.), *The acquisition of Spanish morphosyntax: The L1/L2 connection* (pp. 153–178). Dordrecht: Kluwer.

Cartwright, T. A., & Brent, M. R. (1997). Syntactic categorization in early language acquisition: Formalizing the role of distributional analysis. *Cognition*, *63*, 121–170.

Davies, M. (2006). *A frequency dictionary of Spanish: Core vocabulary for learners*. London: Routledge.

Dixon, R. M. W. (1982). *Where have all the adjectives gone?* Berlin: Mouton de Gruyter.

Durieux, G., & Gillis, S. (2001). Predicting grammatical classes from phonological cues: An empirical test. In B. Höhle & J. Weissenborn (Eds.), *Approaches to bootstrapping: Phonological, syntactic and neurophysiological aspects of early language acquisition* (pp. 189–229). Amsterdam: Benjamins.

Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). New York: Cambridge University Press.

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 63–103). Oxford: Blackwell.

Finch, S., & Chater, N. (1992). Bootstrapping syntactic categories. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society of America* (pp. 820–825). Mahwah, NJ: Erlbaum.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Finch, S., & Chater, N. (1994). Learning syntactic categories: A statistical approach. In M. Oaksford & G. D. A. Brown (Eds.), *Neurodynamics and psychology* (pp. 249–321). San Diego, CA: Academic Press.

Gerken, L. A. (2001). Signal to syntax: Building a bridge. In B. Höhle & J. Weissenborn (Eds.), *Approaches to bootstrapping: Phonological, syntactic and neurophysiological aspects of early language acquisition* (pp. 147–166). Amsterdam: Benjamins.

Gerken, L. A., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, *32,* 249–268.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1,* 3–55.

Golinkoff, R. M., Hirsh-Pasek, K., & Schweisguth, M. A. (2001). A reappraisal of young children's knowledge of grammatical morphemes. In B. Höhle & J. Weissenborn (Eds.), *Approaches to bootstrapping: Phonological, syntactic and neurophysiological aspects of early language acquisition* (pp. 167–188). Amsterdam: Benjamins.

Grimshaw, J. (1981). Form, function, and the language acquisition device. In C. L. Baker & J. J. McCarthy (Eds.), *The logical problem of language acquisition* (pp. 163–182). Cambridge, MA: MIT Press.

Hakuta, K. (1987). The second-language learner in the context of the study of language acquisition. In P. Homet, M. Palij, & D. Aaronson (Eds.), *Childhood bilingualism: Aspects of linguistic, cognitive, and social development* (pp. 31–55). Mahwah, NJ: Erlbaum.

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

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. New York: Longman.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25,* 603–634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language acquisition. *Language Learning*, *57,* 1–33.

Kelly, M. (1996). The role of phonology in grammatical category assignment. In J. Morgan & K. Demuth (Eds.), *From signal to syntax* (pp. 249–262). Mahwah, NJ: Erlbaum.

Labelle, M. (2005). The acquisition of grammatical categories: The state of the art. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 433–457). Amsterdam: Elsevier.

Lafford, B. A., & Collentine, J. G. (1987). Lexical and grammatical access errors in the speech of intermediate/advanced level students of Spanish. *Lenguas Modernas*, *14,* 87–112.

Lafford, B. A., & Collentine, J. G. (1989). The telltale targets: An analysis of access errors in the speech of intermediate students of Spanish. *Lenguas Modernas*, *16*, 143–162.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Liceras, J. M., & Diaz, L. (2000). Triggers in L2 acquisition: The case of Spanish N-N compounds. *Studia Linguistica*, *54,* 197–211.

Maratsos, M. P. (1988). The acquisition of formal word classes. In Y. Levy, I. M. Schlesinger, & M. Braine (Eds.), *Categories and processes in language acquisition* (pp. 31–44). Mahwah, NJ: Erlbaum.

Maratsos, M. P., & Chalkley, M. A. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language* (Vol. 2, pp. 127–214). New York: Gardner Press.

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory and Cognition*, *30,* 678–686.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90,* 91–117.

Mintz, T. H. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 31–63). New York: Oxford University Press.

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, *26,* 393–425.

Morin, R. (2003). Derivational morphological analysis as a strategy for vocabulary acquisition in Spanish. *Modern Language Journal*, *87,* 200–221.

Morin, R. (2006). Building depth of Spanish L2 vocabulary by building and using word families. *Hispania*, *89,* 170–182.

Odlin, T., & Natalicio, D. (1982). Some characteristics of word classification in a second language. *Modern Language Journal*, *66,* 34–38.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Piera, C. (1995). On compounding in English and Spanish. In H. Campos & P. Kempchinsky (Eds.), *Evolution and revolution in linguistic theory* (pp. 302–315). Washington, DC: Georgetown University Press.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Redington, M., & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: A distributional perspective. *Language and Cognitive Processes, 13,* 129–191.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science, 22,* 425–469.

Scarcella, R., & Zimmerman, C. B. (1998). Academic words and gender: ESL student performance on a test of academic lexicon. *Studies in Second Language Acquisition*, *20,* 27–49.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, *19*, 17–35.

Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, *36,* 145–171.

Slabakova, R. (2002). The compounding parameter in second language acquisition. *Studies in Second Language Acquisition*, *24,* 507–540.

Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, *28,* 387–422.

Terrell, T., Andrade, M., Egasse, J., & Muños, E. M. (2005). *Dos Mundos* (6th ed.). Boston: McGraw Hill.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language*, *28,* 649–667.

White, L., Valenzuela, E., Kozlowska-Macgregor, M., & Leung, I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, *25,* 105–133.

Whitley, S. (2002). *Spanish/English contrasts*. Washington, DC: Georgetown University Press.

Whitley, S. (2004). Lexical errors and the acquisition of derivational morphology in Spanish. *Hispania*, *87,* 163–172.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

APPENDIX A

CODING GUIDELINES

Word class errors include the use of a target language word in an inappropriate grammatical frame.These types of errors were excluded from our analysis:

1. Due to transfer of the gerund from English
   Example: *Malos modales son hablando con comida en su boca*. [hablar]
   "Bad manners are talking with food in your mouth."
2. Due to the use of non-Spanish forms
   Example: *Cuando una persona abierte la puerte a mí* [abre]
   "When a person opens the door for me."
3. Stemming from the omission of words
   Example: *Posiblemente yo digé alguna mala sobre…* [alguna mala cosa]
   "Possibly I said some bad about…"
4. Due to confusion between *bien–bueno* "well–good" and *mal–malo* "badly–bad"
   Example: *Para mí ser un bien ejemplo* [buen]
   "For me to be a well example"
5. Due to the use of an infinitive in place of a conjugated verb
   Example: *Hay muchos razones que la gente mentir* [miente]
   "There are many reasons that people to lie"

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

APPENDIX B

WORD CLASS ERRORS PRODUCED IN WRITTEN TASKS

Type	Number of tokens	Example
N for an Adj	8	Los padres son *responsabilidades* [responsables] para el problema "The parents are responsibilities for the problem"
N for an Adv	1	*Especialidad* [especialmente] sus abuelos "Specialty their grandparents"
N for a V	3	Muchos modelos de familias *funcion* [funcionan] "Many models of families function"
N for an Inf	2	Yo no pienso aceptable a *mentira* [mentir] "I don't think acceptable to lie"
V for an Adj	16	Es *necesito* [necesario] mentir cuando "It's I need to lie when"
V for a N	38	Me dijo una *mienta* [mentira] "He told me a lie"
V for an Adv	1	*Incluye* [incluso] las tiendas "Includes the stores"
V for an Inf	23	Buenos modales son *escucha* [escuchar] a sus padres "Good manners are listen to your parents"
Adj for a N	24	Beber mucho *alcoholica* [alcohol] con frecuencia "To drink a lot of alcoholic often"
Adj for a V	5	Yo *necesario* [necesito] las buenas notas "I necessary good grades"
Adj for an Adv	3	Me gusta mi familia *mucha* [mucho] "I like my family a lot"
Inf for a N	3	Los padres tienen mas *preocuparse* [preocupaciones] "The parents have more to worry"
Inf for an Adj	4	Comer con su boca *abrir* [abierta] "To eat with your mouth to open"
Adv for an Adj	1	Produce niños tan *normalmente* [normales] que "Produce children so normally that"

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

APPENDIX C

TARGET ITEMS INCLUDED IN THE FCT

Verbs Used as Adjectives (*k* = 8)

1. Mis padres no me permitieron comer con la boca (abra / abierta).
2. Mi padre dice que es (necesario / necesito) que saques la basura.
3. En un restaurante elegante es malo si no comes con la boca (cierra / cerrada).
4. A veces las mentiras pequeñas son (válidas / vale).
5. No es (importa / importante) que los niños vivan con sus abuelos.
6. Los hijos están (activos / activan) en deportes.
7. También es (divierte / divertido) jugar con los primos.
8. Los deportes tienen muchas personas que son muy (importantes / importan) para los niños.

Verbs Used as Nouns (*k* = 8)

1. Es un problema cuando hay mala (intenta / intención).
2. Los hijos necesitan (disciplina / disciplino) y mucha amistad también.
3. Las (pelean / peleas) causan problemas dentro de la familia.
4. Una persona miente sobre muchas cosas: su (trabajo / trabaja), su familia, y su pasado.
5. Las personas dicen (mientes / mentiras) porque tienen miedo.
6. Las (mentiras / mientes) ligeras están bien porque hacen feliz a la persona.
7. Es aceptable mentir cuando no quieres afectar los (sientes / sentimientos) de una persona.
8. Si a la persona no le gusta tu (respuesta / responde), va a preguntar otra vez.

Verbs Used as Infinitives (*k* = 8)

1. Yo creo que es importante para los hijos (tienen / tener) contacto con sus parientes.
2. Es bueno para los abuelos, hijos, y tíos (pasar / pasan) tiempo juntos.
3. Un ejemplo de buenos modales es (tiene / tener) el respeto por otras personas.
4. Ejemplos de buenos modales son (abrir / abre) la puerta para otras personas.
5. Es aceptable (miente / mentir) cuando la verdad no es simpática.
6. Es más difícil (decir / dice) la verdad pero es lo mejor.
7. Creo que (miente / mentir) está mal.
8. Mis padres me enseñaron a (comer / come) todo en mi plato.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press

Adjectives Used as Nouns (*k* = 8)

1. Cuando muchas personas reciben educación, es mejor para (el económico / la economía) del país.
2. El (divorcio / divorciado) puede causar problemas dentro de una familia.
3. Una familia feliz consiste en mucho amor y (cariñosa / cariño) en una casa.
4. No es necesario vivir con la mamá y el papá para tener (felicidad / feliz).
5. Una razón para mentir es proteger (los emocionados / las emociones) de un amigo.
6. Cuando una amiga tiene un (corte / corto) de pelo muy feo es aceptable mentir.
7. En otras situaciones mentir ayuda a la gente a ganar una posición de (fuerte / fuerza).
8. Una familia feliz consiste en amor, (comodidad / cómodo), y comida.

https://doi.org/10.1017/S0272263109090019 Published online by Cambridge University Press