

Input from Spanish Textbooks: Two Case Studies of Poverty/Richness of the Stimulus

Grant Goodall

University of California, San Diego

The “logical problem of language acquisition” refers to the fact that the evidence available to the child appears to underdetermine the knowledge of language that is ultimately attained (e.g., Wexler and Culicover (1980)). There are two traditional research strategies that try to address this problem. One explores the idea that the environment may in fact be richer than it appears at first (e.g., Tomasello (2003)), and the other pursues the idea that it is the child’s innate language-specific structures that are richer than commonly thought (e.g., Pinker (1995)).

In both of these research traditions, the focus has mostly been on child acquisition of a first language, but there is evidence that the logical problem of language acquisition arises in the case of adult second language acquisition as well (e.g., Schwartz and Sprouse (2000), White (2003)). This case is messier in some ways, in that adults bring a wider range of prior experience to the task than do children, and the adult’s environment may be more varied as well. When the adult’s environment consists of instructional (classroom) settings, there may be additional factors that complicate the acquisition picture. One such factor arises from the simple fact that in the typical classroom setting, the learner is placed in an environment where s/he is surrounded by people who don’t know the language in question (or at best, have incomplete knowledge of it). There are thus some inherent difficulties in trying to create an appropriate linguistic environment in such circumstances.

Despite these inherent difficulties, there are some aspects of the instructional environment that may be advantageous to acquisition. First, as has often been noted, the instructional environment is conceivably richer than a “natural” environment in a number of ways. For instance, instructors may provide explicit negative evidence, which would not otherwise be readily available in the environment. Second, the instructional environment may be impoverished in a way that could facilitate adult acquisition. For instance, if the declining ability of adults to acquire a second language is partly due to the fact that their working memory capacity allows them to process relatively large chunks of language, which in turn allows them to ignore crucial details in the input (Elman (1993), Newport (1988, 1990)), it is at least imaginable that instructional input would consist largely of smaller chunks, in which it would be less possible for adults to ignore these details.

In this paper, I explore two types of instructional input that plausibly have the opposite effect, i.e. that would appear to be disadvantageous to acquisition. Both types involve classroom acquisition of Spanish as a second language. The first type concerns progressive verb forms, a case where the classroom environment appears to be richer than what would be available otherwise. The second type concerns pronominal verbs, where the classroom environment is very impoverished. Both of these cases will involve a comparison between the input available in the instructional environment and that which would be available in a naturalistic setting, such as that to which an adult would be exposed in the course of everyday life. For both of these types of input, I will use methods to approximate the relevant details. For instructional input, I will analyze the way in which grammatical structures are used in three of the most widely used textbooks for Spanish as a second language in North America. As Anderson (2007) has demonstrated, live classroom input seems to follow textbook organization very closely, so this analysis should give us a close approximation of what classroom input is like in many cases. For natural input, I will rely on the Corpus del Español (Davies (2002-)), a tagged corpus that has the capability of being restricted to particular genres or historical periods. For this study, I will use oral texts from the 20th century as an approximation of the environment available to learners in natural settings.

The main focus throughout this study will be on the input available to the learner and on the potential implications of the nature of this input for acquisition, rather than on the process of acquisition *per se*.

1. Case #1: Progressive verb forms

Progressive verb forms in Spanish have the morphological structure in (1), and are often accompanied by a finite form of *estar* ‘to be’, as in (2).

(1) root + theme vowel + *ndo*

(2) *Estoy hablando*
 be.1s speak-PROG
 ‘I am speaking.’

In order to compare the frequency of the progressive relative to other verb forms in Spanish, a count was conducted of verb forms in 20th-century oral texts in the Corpus del Español, with results as shown in Table 1 (progressive verb form represented as “-NDO”; past participle represented as “-DO”).

verb form	count	% of total
present	376,250	47.8
infinitive	143,839	18.3
-DO	63,829	8.1
imperfect	57,134	7.3
preterite	54,079	6.9
subjunctive - present	38,228	4.9
-NDO	24,657	3.1
conditional	10,526	1.3
future	9,647	1.2
subjunctive - ra	7,596	1.0
subjunctive - se	753	0.1
subjunctive - future	34	0
TOTAL	786,572	100

Table 1: Frequency of verb forms in 20th-century oral texts in Corpus del Español

Out of the 786,572 verb forms in this section of the corpus, only 24,657 (3.1%) were in the progressive. In natural input, then, it appears that this verb form is relatively infrequent.

We now turn to instructional input, by means of an analysis of the progressive in common textbooks. One obvious way in which textbooks differ from natural input is that in the former, the introduction of grammatical structures is staggered, with the result that for a significant period of time, certain forms may be virtually absent from the input, while those that have already been introduced dominate. In general, then, the earlier a particular form is introduced, the higher its frequency in the overall input. New forms will be very frequent in the textbook chapter in which they are introduced, and they are then eligible for use in any subsequent chapter (and since “review” is built into most textbook chapters, it is virtually certain that the form will appear again in subsequent chapters).

The point at which a structure is used for the first time thus gives us a rough indication of its frequency in the overall input. The chapter in which the progressive is introduced in three widely-used

textbooks is shown in Table 2 (A = Knorre, Dorwick, Pérez Gironés, and Glass (2008), B = Terrell, Andrade, Egasse, and Muñoz (2005), and C = Blanco and Donley (2007)).

Textbook	Chapter in which progressive is introduced	Total number of chapters
A	5	18
B	5	16
C	5	18

Table 2: Point at which progressive is introduced in three widely-used textbooks

The three textbooks are remarkably uniform, in that in each case the progressive is introduced within the first third of the book, which suggests that its frequency in the input will be relatively high.

Given what we have seen so far, there appears to be an interesting difference between natural input and the instructional input we have examined here, in that roughly speaking, the progressive form is relatively infrequent in the former but relatively frequent in the latter. We will examine the nature of this difference more carefully as we proceed, but let us now turn to the question of whether a difference of this type would have consequences for the acquisition of the progressive. To do this, let us consider separately the morphological form of the progressive and its meaning, since these are the two major facets of it that must be learned.

Morphologically, the progressive has a very simple structure, as we saw in (1), and it is highly regular (a few verbs exhibit a semi-regular stem vowel alternation, but everything else is fully regular). Acquiring the morphology of this verb form thus requires exposure to a range of verbs (minimally, representatives of the *-ar* and *-er/-ir* classes), but because of its simplicity and regularity, relatively little exposure should be sufficient.

With regard to meaning, there does not appear to be much that an L1 English speaker would need to learn, given the close similarity in the meaning of the progressive in the two languages. Nonetheless, learners must figure out how the progressive fits into the overall tense/aspect system of Spanish. In particular, they must acquire the fact that the simple present in Spanish, unlike in English, is non-perfective (Cowper 2005), with the result that Spanish (3a) is possible, unlike English (3b).

- (3) a. *¿Adónde vas?*
 where go.2s
 ‘Where are you going?’
- b. **Where do you go?* (in a non-habitual reading)

Examples such as (3a) are presumably very plentiful in the input (and in environments where the intended meaning would be very clear), so one would not expect this non-perfective nature of the present in Spanish to be difficult to acquire. Furthermore, under the assumption that simpler forms tend to block more complex forms, the learner would correctly conclude that the existence of (3a) means that progressive forms like (4a) must be marked, and thus used primarily in cases where one has some reason to emphasize the progressive nature of the event.

- (4) a. *¿Adónde estás yendo?*
 where be.2s go-NDO
 ‘Where are you going?’
- b. *Where are you going?*

In English, on the other hand, the non-existence of (3b) allows (4b) to be an unmarked form, i.e., the primary way to express the progressive nature of a (non-stative) event. Knowledge of the marked status of (4a) in Spanish thus follows from the possibility of a progressive reading for (3a), but it should also follow from the simple fact that one hears forms such as (4a) very infrequently. In an

important sense, then, one acquires the meaning of the progressive form in Spanish through exposure to many examples like (3a) and few examples like (4a).¹

This latter idea, that infrequency of exposure helps acquisition of the progressive, is reminiscent of the situation facing learners acquiring the use of overt subject pronouns in Spanish. From an acquisition perspective, the fact that the overt pronoun is the marked (and thus emphatic or contrastive) option follows from two things: the existence of a simpler, unmarked option (the null pronoun) and the relative infrequency of overt pronouns. Here too, then, the lack of exposure to the form in question would be expected to play a role in making the use of this form acquirable.

Taking into consideration now the acquisition of both the morphology and the meaning of the progressive in Spanish, we see that neither would appear to require large amounts of input. With regard to the morphology, relatively little exposure is sufficient, given the simplicity and regularity of the form, and with regard to the meaning, large amounts of exposure to the progressive would in fact be unhelpful, since that would mask the marked nature of this form.

With these considerations in mind, let us now return to the two types of input discussed above. In natural input, as approximated by the corpus study summarized in Table 1, we saw that only 3.1% of verb forms were progressive. Given what we have just seen about acquisition, this low frequency would appear to be sufficient for acquiring the morphology, and useful for acquiring the meaning. This is perhaps not a surprising result, since although we have been discussing L1 English speakers acquiring L2 Spanish, very similar factors would presumably come into play in child L1 acquisition of Spanish, and we would expect that naturally occurring input would facilitate this acquisition.

The characteristics of instructional input, as approximated by the analysis of textbooks, were very different, since we saw there that the progressive form is introduced early, with the result that the frequency of the form will be relatively high. It is not practical to obtain an exact measure of the frequency, but we can approximate it by assuming that when a new verbal form is first introduced, it constitutes 40% of the input in that chapter, and that in subsequent chapters, that form will form part of the remaining 60%.

The results of making these assumptions, together with the timing of the introduction of new verb forms in the three textbooks examined, are presented in Figure 1 (notice that the progressive is introduced in chapter 5, as we saw in Table 2 above, and the preterite in chapter 6). There are two points of interest in Figure 1. First, exposure to the simple present diminishes with time, as other verb forms are introduced. By week 5, it has dropped to a level significantly below the 47.8% seen in the corpus study (see Table 1). Second, the progressive constitutes a very significant portion of the input (20–40%) for a considerable period of time (six weeks, if one follows the standard pattern of devoting two weeks to each textbook chapter). Notice that one can adjust significantly the assumptions made in producing this approximation of input in Figure 1 (by positing that new forms constitute more or less than the 40% assumed here, for instance), but these two points will still hold in rough outline.

The extended, large amount of exposure to the progressive should be helpful for acquiring the morphological form, in that learners would have access to abundant evidence for this. As we have seen, however, such massive exposure would not seem to be necessary, at least relative to other verb forms, given the simplicity and regularity of the progressive. With regard to acquisition of the meaning of the progressive, the implications of Figure 1 are not so benign. The overrepresentation of the progressive in the input for an extended period of time, together with the increasing underrepresentation of the simple present, could reasonably be expected to work against the conclusion that the learner eventually needs to reach: that the progressive is a marked form in Spanish.

¹ Another fact that L1 English speakers have to acquire is that an imminent-event reading is disallowed in the Spanish progressive, as seen in (ia) (cf. English (ib)).

- (i) a. **Me estoy juntando con Julie en cinco minutos.*
 1s be.1s join-NDO with in five minutes
 'I'm getting together with Julie in five minutes.'
- b. *I'm getting together with Julie in five minutes.*

I will not pursue this interesting puzzle here.

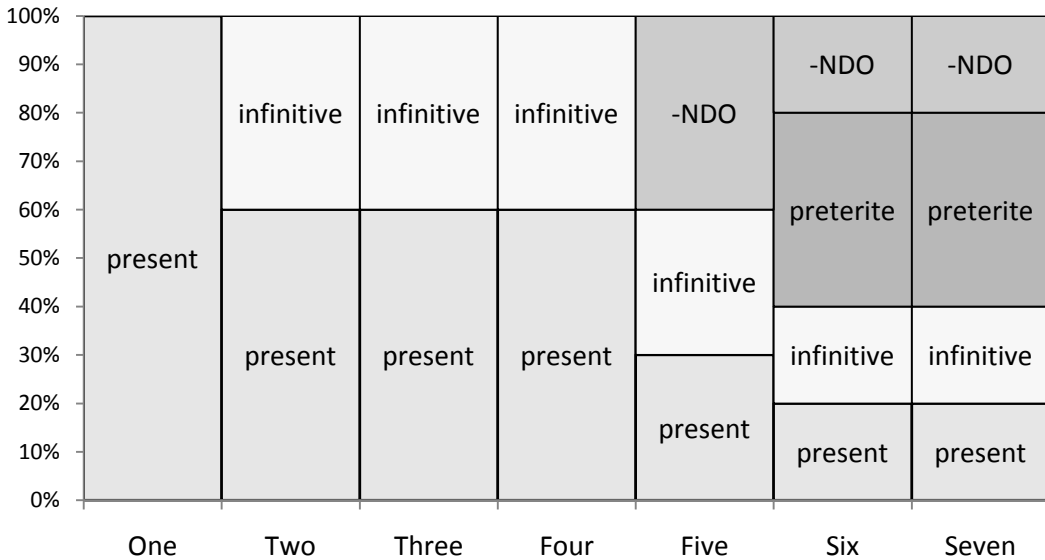


Figure 1: Percentage of input of four verb forms in early textbook chapters

2. Case #2: Pronominal verbs

We now turn to pronominal verbs in Spanish, another case in which we will find interesting differences between the input in natural and instructional contexts. I use the term “pronominal verb” to refer to any verb form in which the inflectional ending on the verb itself agrees in person and number with an object clitic pronoun, as in (5):

- (5) *Me acuerdo.*
 1s remember.1s
 ‘I remember’

Pronominal verb morphology as in (5) is used in a variety of ways in the language. The most straightforward is that in which the clitic pronoun is understood reflexively, i.e. more or less literally as the object of the verb. Examples of this are given in (6).

- (6) a. *Me veo*
 1s see.1s
 ‘I see myself’
- b. *Me compro...*
 1s buy.1s
 ‘I buy ... (for) myself’
- c. *Me encuentro*
 1s find.1s
 ‘I find myself’
- d. *Me considero...*
 1s consider.1s
 ‘I consider myself...’
- e. *Me dedico*
 1s dedicate.1s
 ‘I dedicate myself’

Many other uses of pronominal verbs do not have this reflexive reading. For instance, many verbs of mental processes or states are pronominal, as in (5) above and (7).

- (7) a. *Me imagino*
1s imagine.1s
'I imagine'
- b. *Me alegro*
1s make-happy.1s
'I am happy'

Pronominal morphology may also add aspectual information in a number of cases, as in (8).

- (8) a. *Me comí el pan.*
1s eat.PRET.1S the bread
'I ate the bread (completely).'
- b. *Me fui.*
1s go.PRET.1s
'I departed.'
- c. *Me caí.*
1s fall.PRET.1s
'I fell down.'
- d. *Me dormí.*
1s sleep.PRET.1s
'I fell asleep.'

I will not attempt an exhaustive classification of pronominal verbs in Spanish here (much less an analysis), but I will note that there are many examples that do not fit easily into any of the categories discussed so far:

- (9) a. *Me atrevo a ...*
1s dare.1S to
'I dare to...'
- b. *Me veo bien.*
1s see.1s well
'I look good.'
- c. *Me equivoco.*
1s wrong.1s
'I am wrong.'
- d. *Me enteré.*
1s inform.PRET.1s
'I found out.'
- e. *Me canso.*
1s tire.1s
'I get tired.'

For the purposes of comparing pronominal verbs in natural and instructional input, I will contrast reflexives as in (6) with all other types. I will operationalize the notion “reflexive” as being any pronominal verb that can be naturally translated into English as a verb with a reflexive object. As in the previous section, natural input was approximated by examining 20th-century oral texts in the Corpus del Español. All 1st-person singular verb forms preceded by *me* were extracted (thus avoiding potential complications with 3rd-person *se*, which has non-reflexive uses), and out of these, the 100 most frequent, all containing at least 8 instances each, were retained for further analysis. Cases in which *me* is followed by a clitic-climbing verb, as in (10), were excluded from the analysis, since here the clitic does not indicate pronominal status for the 1st-person verb.

- (10) *Me puedo acordar.*
 1s can.1s remember
 ‘I can remember.’

Notice that *me puedo* by itself is not possible, and (10) is thus unambiguously a clitic-climbing structure.

This procedure resulted in 4042 pronominal verbs, which were then manually classified as reflexive or not. Verbs where the preceding clitic pronoun could potentially be the result of clitic-climbing (as opposed to verbs like *puedo* in (10), for which clitic-climbing is the only possibility), were sampled manually to determine the proportion of reflexives versus other types of pronominal verbs. For instance, the verb *ir* ‘go’ with a preceding clitic pronoun may be the result of clitic-climbing, as in (11), or not, as in (12).

- (11)a. *Me voy a dedicar...*
 1s go.1s to dedicate
 ‘I am going to dedicate myself...’
- b. *Me voy a acordar.*
 1s go.1s to remember
 ‘I am going to remember.’

- (12) *Me voy.*
 1s go.1s
 ‘I am departing.’

Examples such as (11a) were counted as reflexives, but those such as (11b) or (12) were not.

The results of the classification of pronominal verbs as reflexives or not are given in Table 3.

	Count of tokens	% of total
Reflexives	468	12
Other pronominal verbs	3574	88
Total	4042	100

Table 3: Reflexives versus others in 20th-century oral texts in *Corpus del Español*

As Table 3 shows, the reflexives were a relatively small portion of the overall number of pronominal verbs. Part of the reason for this is that the most frequent pronominal verbs are generally not reflexives. Table 4, for instance, shows that of the ten most common pronominal verbs in the corpus, only one, *encontrarse*, had a significant number of reflexives.²

² The verbs are given here in their citation infinitive form, but the counts reflect verb forms with 1st-person singular agreement, regardless of tense/aspect/mood inflection.

Verb	#
acordarse ‘remember’	735
irse ‘depart’	610
imaginarse ‘imagine’	301
sentirse ‘feel’	228
referirse ‘refer’	181
quedarse ‘stay behind’	172
darse (idiomatic)	162
encontrarse ‘find’	146
ponerse ‘start’	140
casarse ‘get married’	65

Table 4: Ten most frequent pronominal verbs in 20th-century oral texts in *Corpus del Español*

Sampling of the tokens for *encontrarse* indicate that they are almost evenly divided between a reflexive reading (e.g., *me encontré en la plaza* ‘I found myself in the plaza’) and other readings (e.g., *me encontré con alguien* ‘I ran into someone’ or *me encontré con que...* ‘I discovered that...’). Two of the other verbs, *darse* and *ponerse*, do allow a reflexive reading in principle (e.g., *me di dinero* ‘I gave myself money’ or *me puse el pantalón* ‘I put the pants on (myself)’), but these readings were not found in the sampling of the tokens of those verbs, where the overwhelming majority of the examples were of the type *me doy cuenta* ‘I realize’, *me di la vuelta* ‘I turned around’, or *me puse a...* ‘I started to...’. The relatively small number of reflexives overall is thus reflected in this list of the most frequent pronominal verbs.

Now that we have a characterization of the input that would be likely in a natural setting, as revealed by the corpus analysis, let us turn to an examination of the input in an instructional setting, by means of an analysis of pronominal verbs in the three textbooks mentioned earlier. If we look at which pronominal verbs are used in these three textbooks, we find similar results across all three: reflexives predominate, and other pronominal verbs are given scant attention, if any.³ These results are summarized in Table 5.

Textbook	% reflexives	# of verbs from Table 4
A	90	0
B	93	0
C	82	2

Table 5: Reflexives and most frequent pronominal verbs in three textbooks

As Table 5 shows, the great majority of the pronominal verbs introduced are reflexives, and the most common pronominal verbs (few of which are reflexives, as we have seen) are generally not introduced at all.

We thus see a stark difference between natural and instructional input with regard to the proportions of reflexives versus other pronominal verbs. Reflexives are just one part of the overall pronominal verb input in natural settings, whereas they form almost all of the pronominal verb input in instructional settings. What are the expected consequences of this difference for acquisition? One obvious consequence is that learners exposed only to instructional input would seem to be at a considerable disadvantage, in that many high-frequency pronominal verbs are simply absent from the input and thus not able to be acquired.

There are also less obvious ways, however, in which the instructional input could impair acquisition. Under some proposals for pronominal verbs in Spanish (e.g., de Miguel and Fernández Lagunilla (2000)), the various descriptive subtypes of pronominal verbs, some of which we sampled

³ A few other pronominal verbs do show up in isolated instances in the later chapters in each book, but the count in Table 5 reflects those pronominal verbs for which there is substantial input (i.e. more than a single mention).

above, are argued to have a single, unified analysis. If such an approach is correct, then the diverse range of pronominal verb types available in natural input would seem to support convergence by learners on this sort of analysis. With the instructional input, on the other hand, learners would have only limited access to a significant part of the relevant data, which would make converging on this type of analysis extremely difficult, if not impossible.

In addition, the absence of many of the highest-frequency pronominal verbs in textbooks could conceivably result in fewer pronominal verbs overall, since one would presumably use non-pronominal synonyms of these verbs. This in turn could make acquisition of the morphological form of pronominal verbs more difficult. Notice that the form of pronominal verbs is relatively complex (unlike what we saw earlier with progressive verb forms, for example). Agreement with the subject occurs both as an inflectional ending on the verb and as the clitic pronoun, so one would expect a relatively long course of acquisition, and a reduction in the amount of input available would presumably delay acquisition even further.

There thus appear to be a number of reasons why the more concentrated exposure to reflexives in the instructed input could result in impaired acquisition, compared to the more diverse range of pronominal verbs found in natural input.

3. Conclusions

We have now seen evidence that instructional input regarding the two types of verbal morphology in Spanish that have been examined here, the progressive and pronominal verbs, differs in significant ways from natural occurring input, and that these differences could be detrimental to acquisition. In effect, this is because instructional input makes the logical problem of language acquisition worse in these two instances, by widening the gap between the evidence available in the input and what must eventually be acquired. In the case of the progressive form, the gap results paradoxically from an overrepresentation of this form in the input, while in the case of pronominal verbs, it results from an underrepresentation of certain types of pronominal verbs.

Looking more specifically at acquisition of the form (morphology) and meaning, the consequences of the overrepresentation of the progressive in the instructional input should be neutral with regard to the morphology (since acquisition should be possible with relatively small amounts of exposure), but negative with regard to the meaning. The consequences of the underrepresentation of certain types of pronominal verbs, on the other hand, should be negative for both the morphology and the meaning.

It is worth emphasizing that the present study has analyzed the input available for acquisition, in both natural and instructional settings, and not the acquisition process itself in either environment. As we have seen, though, one can draw reasonable conclusions about the consequences for acquisition that could stem from the distinctive properties of instructional input. It remains to be explored how large the detrimental effects are in real-life instructional contexts and how instructional practice can be modified to mitigate them. Even in advance of fully answering these questions, however, we have seen here that there is ample reason for great care in designing input for instructional settings. As in much else, opportunities for unintended consequences abound.

References

- Anderson, Bruce (2007). Pedagogical Rules and their Relationship to Frequency in the Input: Observational and Empirical Data from L2 French. *Applied Linguistics* 28, 2: 286–308.
- Blanco, J. and P. R. Donley (2007). *Vistas*, 3rd edition. Vistas Higher Learning.
- Davies, Mark. (2002-). *Corpus del Español (100 million words, 1200s-1900s)*. Available online at <http://www.corpusdelespanol.org>.
- de Miguel, Elena & M. Fernandez-Lagunilla. (2000). El Operador Aspectual SE. *Revista Española de Linguística* 30, 1: 13-43.
- Elman, Jeff (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, 48: 71-99.
- Knorre, M., T. Dorwick, A. M. Pérez Gironés, and W. R. Glass (2008). *Puntos de Partida*, 8th ed. McGraw-Hill.

- Newport, Elissa (1988). Constraints on learning and their role in language acquisition: Studies of the acquisition of American sign language. *Language Sciences* 10, 1: 147-172.
- Newport, Elissa (1990). Maturational constraints on language learning. *Cognitive Science* 14, 1: 11-28.
- Pinker, Steven (1995). *The language instinct*. Harper Collins.
- Schwartz, Bonnie and Rex Sprouse (2000). When syntactic theories evolve: Consequences for L2 acquisition research. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 156-186). Malden, MA: Blackwell.
- Terrell, T., M. Andrade, J. Egasse, and E. M. Muñoz (2005). *Dos Mundos*, 6th ed. McGraw-Hill.
- Tomasello, Michael (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
- Wexler, Kenneth and Peter Culicover (1980). *Formal principles of language acquisition*. Cambridge, MA: MIT Press.
- White, Lydia (2003). *Second language acquisition and Universal Grammar*. New York: Cambridge University Press.

Selected Proceedings of the 12th Hispanic Linguistics Symposium

edited by Claudia Borgonovo,
Manuel Español-Echevarría,
and Philippe Prévost

Cascadilla Proceedings Project Somerville, MA 2010

Copyright information

Selected Proceedings of the 12th Hispanic Linguistics Symposium
© 2010 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-440-9 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

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

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

Web access and citation information

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

This paper can be cited as:

Goodall, Grant. 2010. Input from Spanish Textbooks: Two Case Studies of Poverty/Richness of the Stimulus. In *Selected Proceedings of the 12th Hispanic Linguistics Symposium*, ed. Claudia Borgonovo, Manuel Español-Echevarría, and Philippe Prévost, 260-269. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2422.