

Back to *Patrás*: A Process of Grammaticicization in a Contact Variety of Spanish

Daniel J. Villa
New Mexico State University

1. Introduction

In a situation of contact between two languages, a common means of understanding what may appear to be a grammatical innovation is to attribute the transference of a structure from one language to the other.¹ Lipski (1985:91) presents this point of view in observing, "...United States Spanish may employ syntactic constructions which directly or indirectly reflect English structures, and which are not normally found in other Spanish dialects removed from active contact with English". As Otheguy (1991:101) points out, such a process has been identified in a variety of ways: loan translation, calque, loanshift and semantic loan. Such is the case of the construction VERB + *PATRÁS* in U.S. Spanish, as in

- (1) *Y lo'o nos FUIMOS PATRÁS, y duramos cinco años allá....*
And then we WENT BACK, and we stayed there five years....

As this particular structure is striking in its similarity to the English construction VERB + BACK, its appearance has been attributed a translation or calque. Scholars such as Lipski (1985), while recognizing the fact that in general terms the supposed calque follows the grammatical constraints of General Spanish, the structure nevertheless must be understood in terms of borrowing. As Otheguy (1991: 103) observes, 'The thorough and carefully nuanced treatment of Lipski (1985:91ff) presents this construction as "an apparent calque on *back*," and places it in the context of "syntactic Anglicisms"'.

Lipski and Otheguy differ on the source of the construction studied here, but both point to contact with English as the SOURCE for the construction. For example, Otheguy (1991:116) concludes that 'these items [such as *patrás*] represent, not diachronic developments in the lexicon or grammar of the Spanish language, but synchronic adaptations to the largely non-Hispanic cultural and conceptual environments of U.S. Latinos'. In this paper I aim to present another view, that the construction verb + *patrás* represents an internal development in Spanish. Thus, its development can be analyzed using theories of grammaticization and corpus linguistics, which include both synchronic and diachronic dimensions.

2. Theoretical considerations

The analyses presented in this paper rest on two complementary theoretical approaches: grammaticization and corpus linguistics, both relatively recent developments in language study. The former is described in general terms in works such as Heine et al 1991, Traugott and Heine 1991 and Bybee et al. 1994, and for Spanish in Villa 1997 and Torres 2000. While scholars do not agree on all aspects of a general theory of grammaticization, there exist a number of concepts important for this paper that are less polemical than others. One is that the development of a particular grammatical element can be traced over a period of time, and that it follows certain patterns in the changes it undergoes. A second is phonetic erosion or reduction: as a form grammaticizes, it loses phonetic substance. A third is a loss, or bleaching, of semantic content; a semantically rich form becomes one relatively generalized and non-specific in meaning. A fourth is the fixing of the position of the grammatical form to an accompanying element; an item that was relatively syntactically unbound becomes more bound as it grammaticizes.

So, for example, in the development of Spanish synthetic future *-ré* from Vulgar Latin, the verb *habeo* 'I have' phonetically reduces to the first person future marker *-é*, and the semantic content of possession is lost. At one point in the process of grammaticization of this form it was possible to insert

elements between the verb and *-é* (in the following, orthographically *he*), such that in one of the first texts written in modern Spanish, the *Poema de Mio Cid* we find the structure *si yo vivo DOBLAR vos HE la soldada*, ‘if I live I will double your pay’. At that particular juncture in the development of Spanish, it was possible to insert the indirect object pronoun *vos* between the verb *doblar* and the incipient future marker *he*. This is no longer a possibility in modern Spanish, and the orthographic form of the synthetic future is written as a single form, in this instance *doblaré*. I will refer to these four precepts in the analyses below.

The brief example above introduces another important theoretical dimension for this paper, that of the distinction between synchronic and diachronic studies. Christie (1982) introduces the notion of ‘panchronic’ linguistics, a theoretical viewpoint that does not distinguish between synchronic and diachronic studies; he asserts:

In defining panchrony I am not referring to that study of the same name that was discussed by de Saussure and Hjelmslev.... Both of them refer by this term to the study of what we would today refer to as synchronic universals, those principles that will always be and everywhere valid in human languages studied synchronically. Such a study would certainly be part of the panchronic approach, but it certainly would not constitute the whole of the approach. The panchronic linguist will range as widely as the historian, drawing on any piece of information that might illuminate the nature of his subject matter (7).

Within the context of a theory of grammaticization, Hopper (1987) further elaborates on this idea in the framework of what he terms ‘Emergent Grammar’. He writes, ‘Because grammar is emergent but never present, it could be said that it never exists as such, but always coming into being. There is, in other words, no “grammar” but “grammaticization” – movements toward structure which are often characterizable in typical ways’ (10). The importance of these theoretical angles for the present study lies in the fact that, at least in the tradition of Spanish linguistics, forms and structures found in contact varieties of the language are accepted as internal or external based on comparisons with some undefined ‘standard’ Spanish, often a monolingual variety spoken in the capitals of predominantly Spanish speaking countries. A panchronic approach, couched in a theory of grammaticization, seeks to avoid such comparisons, but rather aims to trace the development of a certain structure, in this case verb + *patrás*, based on attested forms deriving from both written and oral texts. This helps to disentangle those phenomena which are contact induced, and those which form a part of the grammaticization process in General Spanish.

This assertion leads into an examination of the relationship between a theory of grammaticization and corpus linguistics. The use of the latter in this paper represents a relatively recent development, particularly in the study of U.S. Spanish, and consequently such an approach does not possess a well developed literature. As Kurtböke (1998:ch.4 para.1) notes,

While various computational approaches have been adopted in language contact before (see Poplack and Sankoff and Joshi among others), the focus has remained traditionally on sentence analysis and with the narrow focus of distinguishing between borrowing and codeswitching and identifying nonce loans. Although corpus-based language contact studies can be encountered today, in particular in the area of historical linguistics, such studies are still few.

With regard to the importance of corpus linguistics for linguistic analyses, Hallan (2001) asserts

The recognition that human beings are embodied communicators, with all that this implies for the human communication system, and the application of psychological findings to explanations in linguistics seem to me to be valuable and constructive. However, the essentially introspection-driven nature of many analyses [of certain scholars] diminishes the value of their theses. ...Their applications... depend on the introspections of individual native-speakers, who for all their extensive training and preminent status in the field, should be exempt from the basic requirement of scientific endeavor, that they find a way of testing their hypotheses. ...[T]he techniques of modern corpus linguistics now make it possible to test at least some of the hypotheses of these [extant] models... (92-3).

I quote Hallan at length as this situation is particularly pertinent in Spanish linguistics in general, and in the study of U.S. Spanish in particular. Both Lipski (1985) and Otheguy (1991) present samples of the use of *patrás* in their articles, yet we have no idea of the frequency of usage or the history of such structures. This is not meant as criticism of those analyses, however, but reflects the reality that large bodies of data, easily searchable, did not exist at the time of writing of those essays. A major stumbling block, at least in a panchronic analysis of Spanish, has been a dearth of electronic corpora searchable by computer.

Fortunately for the field in general and the present work in particular, however, an online corpus recently became available. It is the *Corpus del español*, hereafter CDE, created by Mark Davies at Illinois State University, available on the Internet at <<http://www.corpusdelespanol.org/>>. This corpus represents a collection of written texts dating from the 13th century to the present, and oral texts from the 20th century, for a total of over one hundred million words (see the website for details on the sources). The second source is a collection of oral texts of New Mexican Spanish, a contact variety of the language. Some data were recorded and transcribed by the author. Others come from the New Mexico/Colorado Spanish Survey, a research program for a linguistic atlas of New Mexican and Southern Colorado Spanish carried out by Garland Bills and Neddy Vigil, University of New Mexico. This represents a much more modest corpus of about twenty thousand words. These data, along with those from traditional sources such as grammars and dictionaries, form the bases for the present study.

Regarding the interaction between studies of grammaticization and corpus linguistics on the Spanish language, both Villa (1997) and Torres (2000) have made use of various corpora in their work. However, as Torres (2000:21) notes, such bodies of work have been relatively small in size. But given very recent developments such as the appearance of the Davies corpus, access to large amounts of texts, both written and oral, aids tremendously in tracing the grammaticization of a particular form. Further, the fact that Spanish and Latin have a long history of written texts makes it possible to follow the grammaticization of a particular morpheme for about two millennia.

3. Development of the form *patrás*

Lipski (1985) represents the phrase studied here as *pa(ra) atrás*; Otheguy employs the form *para atrás*, both reflecting analytic forms of this structure. However, the first point I examine is the development of the grammatical morpheme *patrás*. I will argue that the synthetic form studied in this paper is not in some type of distributive relationship with older variants such as *para atrás*, *para tras*, or *pa tras*, but rather a marker in the process of grammaticizing. Toward this end, in the following discussion data from the Davies corpus will be marked CDE; those collected in New Mexico will be tagged with NMS. The search engine available for the CDE was used to search for the variants *para atrás*, *para tras*, *pa atrás*, and *patrás*. After being sorted by century, the data were analyzed for the relationship of the various forms of *patrás* and the semantic content and degree of fusion of the nascent grammatical marker to the accompanying verb form. A word processor search function was used to analyze the NMS data, which were then sorted by degree of grammaticization and type of co-occurring verb.

As noted above, grammatical markers undergo a semantic bleaching, that is, a loss of semantic content. In this particular case, the semantic content lost not only in this form but in other Spanish prepositions as well is the notion of physical motion. According to García de Diego (1985), the Spanish preposition *para* results from the fusion of the Latin prepositions *per* and *ad* (297) and *atrás* from *ad* and *trans* (442). Simpson (1968) notes that in Latin *ad* expresses ‘primarily direction toward’ (10), *per* ‘through, along, over’ (432), and *trans* ‘over, across, on or to the other side of’ (610). As this semantic content of physical movement bleached out in the Latin prepositions they fused, resulting in the creation of new Spanish prepositions, *para* and *atrás* in this case. As can be seen, the loss of semantic content was accompanied by the loss of phonetic substance. Modern dictionaries commonly gloss *para* and *atrás* as prepositions, respectively ‘for’, ‘toward’ or ‘to’ and ‘back’ or ‘behind’.

However, as early as the fifteenth century these ‘new’ prepositions began to collocate in the structure *para atrás* or *para tras*, co-occurring with verbs of motion, such as *volver*, ‘to return’. This indicates a grammatical shift from a prepositional function to an adverbial one. We find in texts from that time:

- (2) ...*boluio el conde PARA TRAS*. [CDE, 15th century]
 ...the count came BACK.

One gloss in modern Spanish dictionaries of *volver* is ‘to return’; at the same time, it can also signify ‘to turn bad’ or ‘to turn over’, among other uses. It would appear that at an early date this verb had begun to lose the semantic content of physical motion toward a point, and that *para tras* added that semantic dimension to the co-occurring verb. This is the case in all the data studied here, that the majority of either the synthetic or the analytic forms modify a verb of motion (see Table 1).

Time period	Para tras		Para atrás		Patrás	
	Move	Non	Move	Non	Move	Non
15 th century (CDE)	5	2	0	0	0	0
16 th century (CDE)	16	0	18	2	0	0
17 th century (CDE)	0	0	2	1	0	0
18 th century (CDE)	0	0	4	0	0	0
19 th century (CDE)	0	0	29	9	1	0
20 th century literary (CDE)	3	2	26	8	0	0
20 th century oral (CDE)	0	0	36	14	2	0
20 th century misc. (CDE)	0	0		0	0	0
20 th century oral (NMS)	1	0	5	1	20	2

TABLE 1. *Patrás* co-occurring with verbs of motion or others

The significance of the predominance of co-occurrence with verbs of motion lies in that a theory of grammaticization predicts a generalization of meaning, in this case from the semantic content of physical movement to an abstract one. Further, in the development of *patrás*, it is worthy of note that from the first appearances of the analytic structure that no element interposes between *para* and *atrás*, another indicator of its entering into a process of grammaticization.

At the same, it can be established that this is a relatively ‘new’ grammatical morpheme. It does not appear in any of the texts predating the fifteenth century, such as the *Poema de Mio Cid*. As noted in example (2) above, it was possible to insert an element between *para tras* and the co-occurring verb ‘*boluio*’, in this instance *conde*. This tendency remains to date in New Mexican Spanish:

- (3) *No sólo eso pero alguien PATRÁS vino....* (NMS)
 Not only that but somebody BACK he came....

Here *patrás* occurs before the verb. While this was the only instance found in the 20th century, such a structure indicates that *patrás* is in a relatively early stage of grammaticization as its postposing is not obligatory.

At the same time, it can be noted that *patrás* can co-occur with verbs that do not indicate physical movement toward a point, another hallmark of a process of grammaticization. The corpora offer:

- (4) ...*MIRANDO PARA ATRÁS, lleno de temor....* (CDE, 16th century)
 ...LOOKING BACK, full of fear....

- (5) ...*el que entra ya no MIRA PARA ATRÁS....* (CDE, 19th century)
 ... he who enters no longer LOOKS BACK....

- (6) *Y le pidió gracias porque le HABÍA DADO sus ojos PATRÁS.* (NMS)
 And he thanked [the Virgin Mary] because she HAD GIVEN him his eyes BACK [i.e. returned his eyesight].

- (7) *Era el maestro muchos años PARA ATRÁS.*
He was the teacher many years BACK.

If it the case that the majority of the various forms of *patrás* co-occur with verb of motion, it is also true that it is panchronically generalizing to appear with verbs that do not indicate movement. The relative infrequency of the latter is another indicator the relatively early stage of grammaticization of *patrás*.

It is perhaps such usages that researchers find the most striking. Again, in the corpora studied here very few instances of *patrás* marking abstract movement appear. Anecdotally, the author has heard such utterances as

- (8) *Ahorita te llamo patrás.*
I'll call you right back.
- (9) *Les hablo en español pero me hablan patrás en inglés.*
I speak to them in Spanish but they talk back [answer] in English.
- (10) *Ya quiero mi troquita patrás.*
I want my truck back now.

These forms exist, but have yet to appear in any corpus, rendering difficult their analysis. What may be asserted is that *patrás* appears to be developing into what I will call here a regressative marker. Such a marker indicates principally physical movement from point A to point B and back to point A again, but also is beginning to refer to a non-physical movement A to B to A, such as in returning a phone call or in speaking to someone. Whatever might be the case, though, I hope to have suggested that this process of grammaticization is one internal to the Spanish language, and does not represent a recent borrowing, or calquing, from English.

To sum up, a panchronic analysis of verb + *patrás* suggests that it has developed due to:

- (a) a loss of semantic content of earlier Latin prepositions,
- (b) a loss of phonetic substance of forms that gave rise to it,
- (c) a need for grammatical morpheme to indicate movement toward, and
- (d) a generalization from physical to non-physical movement, indicating the nascent development of an unbound regressative grammatical morpheme.

4. Implications for studies in language contact

The question then arises, if the construction verb + *patrás* is not a borrowing or calque from English, why does it appear in bilingual Spanish/English communities, apparently with more frequency? Lipski (1985) notes that *patrás* is employed in similar ways by Spanish/English bilinguals in diverse regions of the U.S., in Trinidad, Central America and in Gibraltar. This diversity would seem to support an argument for *patrás* as a calque from English. However, I hope to have shown that precisely the same usage can be found in what are presumably non-contact varieties of Spanish, over the past 500 years or so. Undoubtedly there existed contact between English and Spanish during this time period, but it would have been much less intense in the Iberian Peninsula in the 16th century than in New Mexico in the 20th and 21st.

Given the limitations of the data, particularly the size of the NMS corpus, I cannot offer any firm answers. However, given the commonality of *patrás* in Spanish/English bilingual communities, I will advance a hypothesis. Bybee et al. (1994) have demonstrated that grammatical morphemes for tense, aspect and modality in the languages of the world tend to derive from a limited number of lexical sources. For example, both Spanish and English, as well as many other languages, make use of movement verbs, 'ir' and 'be going to' respectively, in order to express futurity. The periphrastic 'ir' future in Spanish, as in '*voy a pensarlo mañana*', 'I'm going to think about it tomorrow', has so generalized worldwide throughout General Spanish that no researcher, to the best knowledge of the author, suggests that the Spanish 'ir' future derives from the English 'be going to'. *Patrás*, while

undoubtedly much less common than the periphrastic future, shares the same status as a form that results from a process of grammaticization in General Spanish.

In spite of this status, it may be that *patrás* has come to serve as an identity marker of U.S. Spanish speakers in particular and perhaps of Spanish/English bilinguals in general. Much of the research literature on U.S. Spanish focuses on so-called 'non-standard' lexical and syntactic features of the varieties of the language spoken in this country. Often, such 'non-standard' items tend to derive from the working classes whose speech varieties make up the source for the majority of U.S. Spanish dialects. Forms such as *asina* (thus or so), *estábanos* (we were), *hablastes* (you spoke) and *traiba* (he, she or you brought), among many others, are common throughout the Spanish speaking world, occurring in many varieties of General Spanish. These forms, though common, are socially marked, and often used in determining the social positions of speakers. Such may well be the case with *patrás*. Though 'native' to Spanish, it might serve as a marker that one has roots in what some term 'low' varieties of the language (cf. Valdés and Geoffrion-Vinci 1998).

Among the many cultural and social artifacts created here by centuries of migration, one is the perceived inherent superiority of certain varieties of Spanish, again, often those spoken in the capitals of other Spanish speaking countries among certain social classes. As a result, an explicit or implicit idea in research regarding U.S. Spanish is that it represents some sort of an impoverished variety of the language (again, see Valdés and Geoffrion-Vinci 1998 as well as Valdés et al. 2003). This does not represent a surprising situation, since Spanish has very much existed as a subordinate language, or L₂, here in the U.S. since the emerging of this nation as a political entity (see, e.g. García 1993). Forms such as *patrás* allow the Spanish speaking world in general and language researchers in particular to draw lines between certain groups of Spanish speakers. The motives for doing so may differ, but the process is quite similar.

Again, corpus linguistics, wedded to other theoretical approaches such as those of grammaticization, present an important set of tools for challenging certain commonly held notions about language use. What we do not currently know is the distribution of the construction verb + *patrás* throughout the Spanish speaking world. It might be reasonable to suspect that it does occur in other geographic regions, given its panchronic presence in the language. This presence will be difficult to establish until sufficiently large corpora become available, especially those of spoken Spanish. As it represents a newly emerging form, it remains to be seen how long before it becomes common in written texts. However, a quick search of the Internet reveals the fact that *patrás* is used in many geographic areas as a regressative, often to mark on a Spanish web page what is the 'back' button on an English one. The fact that mostly 'hip' pages, and not those of law firms, museums, governmental offices and other social institutions, use *patrás* further suggests that this is newly emerging form.

5. Conclusion

We may reasonably assert that *patrás* does not represent a calque from English; beyond that, it is not possible to determine if its usage in U.S. Spanish varies significantly from that in other varieties of the language. While the author does not share Otheguy's view that the construction verb + *patrás* does not possess a diachronic presence, he does feel that its presence in U.S. Spanish, and its subsequent study in the literature, are due more to social factors than to processes such as borrowing or calquing, as Otheguy suggests. Its development follows empirically based studies in grammaticization, and its presence over the last 500 years in attested documents empirically supports the assertion that it represents an internal development in the Spanish language. It is hoped that the use of various theoretical approaches in trying to establish what phenomena are or are not the result of language contact will be fruitful in future studies in the area of language contact.

Note

I am much indebted to Belén López for her assistance in classifying a portion of the New Mexican Spanish data presented here. I would also like to thank Garland Bills and Neddy Vigil for data from the New Mexico/Southern Colorado Linguistic Atlas Survey, as well as Petek Kurtböke for all her work and patience in organizing the session that this paper was presented in. All errors are mine alone.

References

- BYBEE, JOAN L., REVERE PERKINS, and WILLIAM PAGLIUCA. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- CHRISTIE, WILLIAM M. 1982. Synchronic, diachronic and panchronic linguistics. *Papers from the third international conference on historical linguistics*, ed. by J. Peter Maher, Allan R. Bomhard, and E.F. Konrad Koerner, 1-10. Amsterdam: John Benjamins.
- GARCÍA, OFELIA. 1993. From Goya portraits to Goya beans: Elite traditions and popular streams in U.S. Spanish language policy. *Southwest Journal of Linguistics* 12.69-86.
- GARCÍA DE DIEGO, VICENTE. 1985. *Diccionario etimológico español e hispánico*. Madrid: Espasa-Calpe.
- HALLAN, NAOMI. 2001. Paths to prepositions? A corpus-based of the acquisition of a lexico-grammatical category. Frequency and the emergence of linguistic structure, ed. By Joan L. Bybee and Paul Hopper, 91-120. Amsterdam: John Benjamins.
- HEINE, BERND, ULRIKE CLAUDI, and FRIEDERIKE HÜNNEMEYER. 1991. *Grammaticalization: A conceptual framework*. Chicago: The University of Chicago Press.
- JOSHI, ARAVIND K. 1985. Processing of sentences with intra-sentential code-switching. *Natural language parsing: Psycholinguistic, computational and theoretical perspectives*, ed. by David Dowty, Lauri Karttunen and Arnold Zwicky, 190-200. Cambridge: Cambridge University Press.
- KURTBÖKE, PETEK. 1998. *A Corpus-driven study of Turkish-English language contact in Australia*. Clayton, Australia: Monash University dissertation, electronic version.
- LIPSKI, JOHN. 1985. The Construction pa(ra) atrás among Spanish-English bilinguals: Parallel structures and universal patterns. *Revista/Review Interamericana* 25.91-102.
- MENÉNDEZ PIDAL, RAMÓN. 1968. *Poema de mio Cid. Clásicos castellanos*. Madrid: Espasa-Calpe.
- OTHEGUY, RICARDO. 1991. A Reconsideration of the Notion of Loan Translation in the Analysis of U.S. Spanish. *CUNYForum: Papers in Linguistics* 16.101-21.
- POPLACK, SHANA, and DAVID SANKOFF. 1984. Borrowing: The synchrony of integration. *Linguistics* 22.99-135.
- SIMPSON, D. P. 1968. *Cassell's Latin Dictionary*. New York: Macmillan Publishing Company.
- TORRES CACOULOS, RENA. 2000. *Grammaticization, synchronic variation, and language contact: A study of Spanish progressive -ndo constructions*. Amsterdam: John Benjamins.
- TRAUGOTT, ELIZABETH CLOSS and BERND HEINE. 1991. *Approaches to grammaticalization I and II*. Amsterdam: John Benjamins.
- VALDÉS, GUADALUPE, and MICHELLE GEOFFRION-VINCI. 1998. Chicano Spanish: The problem of 'underdeveloped' code in bilingual repertoires. *Modern Language Journal* 82: 473-501.
- VALDÉS, GUADALUPE, SONIA GONZÁLEZ, DANIA LÓPEZ GARCÍA, and PATRICIO MÁRQUEZ. 2003. Language ideology: The case of Spanish in departments of foreign languages. *Anthropology and Education Quarterly* 34.3-26.
- VILLA, DANIEL. 1997. *El desarrollo de futuridad en el español*. México, D.F.: Grupo Editorial Eón.