

# THE LEXICAL-SYNTACTIC PROCESS OF *RECATORIZATION*: THE CASE OF SPANISH *MEDIO*\*

Enrique Pato and David Heap  
University of Western Ontario

## 1. Introduction

According to traditional descriptive grammars, Spanish *medio* ('half') is a variable category. It can be a noun as in (1a), a numeral adjective (quantifier) as in (1b), or an unstressed adverb as in (1c):

- (1) a. *medio de vida*  
'means of support, resource'.
- b. *media docena*  
'half a dozen'.<sup>1</sup>
- c. *estaba medio cansado*  
'he was somewhat tired'.

However, *medio* takes an invariable form as an adverb in standard Spanish, as in (2a). Making the adverb *medio* agree with the adjective it modifies as in (b) is considered nonstandard.

- (2) a. *estaba medio cansada* (standard)  
'she was somewhat tired'.
- b. *estaba \*media cansada* (nonstandard, agrees in the feminine)  
'she was somewhat tired'.

However, numerous cases are attested where the adverb *medio* agrees, as though it were an adjective, with the noun or the adjective which it modifies. This process, known as *re categorization* (Company 1995, Dubinsky & Williams 1995, Demonte 2000, etc.), implies that an element is interpreted as belonging to one category (i.e. adjective) in some varieties and as part of another (here, adverb) in other varieties:

---

\* The authors would like to thank the participants in the panel on "Language Transfer and Change" during the CLA Conference 2005 at the University Western Ontario for comments and ideas regarding this paper, and Ivan Chow of the Theoretical and Applied Linguistics Lab for his help in preparing the manuscript. This research is supported in part by the Social Sciences and Humanities Research Council of Canada (Heap 410-2003-0461) and the Ministerio de Educación y Ciencia of Spain (Pato EX2004/0970).

<sup>1</sup> A further semantic distinction can be made between dividing entities (leaving a fragment of this entity, e.g. *media manzana* 'half an apple') and dividing quantities (produces another quantity, e.g. *media docena* 'half a dozen' i.e. six), cf. Bosque 1999.

- (3) a. *medio hermanos* (invariable) > adverb.  
 b. *medios hermanos* (agrees m. pl.) ‘half brothers’ > adjective.

## 2. The extent of *medio* regategorization in Iberian Spanish (ALPI data)

The ‘adjectivization’ of *medio* is widely considered to be characteristic of Latin American Spanish varieties. It is commonly found in the vernacular speech of areas such as Argentina, Uruguay, Paraguay, Chile, Peru, Ecuador, Panama, Guatemala, Belize, Mexico, Cuba, Dominican Rep. and Puerto Rico (Kany 1945, Moreno de Alba 1993), but it is believed to be rare in Iberian Spanish.



Figure 1: Spanish Speaking Latin America and the Iberian Peninsula.

Agreement in ‘adjectival’ usage of *medio* (or equivalents) is regularly attested in Portuguese and in Galician (*meio/meia*), and has been described in the Spanish spoken by bilinguals in Galicia (Cotarelo 1927, etc.).

In this study we use unpublished dialect data from the *Linguistic Atlas of the Iberian Peninsula* (*Atlas Lingüístico de la Península Ibérica*, see [www.alpi.ca](http://www.alpi.ca)) to examine the precise geographical area of the phenomenon in Peninsular Spanish, as of the date of these surveys (1930s). The data of interest here come from *ALPI: Cuaderno I*, line #406: *Estaba medio dormida*, to which some subjects surveyed responded *Estaba media dormida*.



Figure 2: Distribution of nonstandard *media* in Peninsular Spanish.  
(• = occurrences of the nonstandard agreeing form *media*)

The distribution of the nonstandard variant with *media* by provinces is as follows:

- In the North: The four Galician provinces Coruña (1), Lugo (2), Pontevedra (3), Orense (4), also north-western León (5) and Zamora (6). In some points of Valladolid (7), Salamanca (8) and Cáceres (9).
- In the South: Huelva (10), Sevilla (11), Cádiz (12) and Málaga (13).
- Also in some other areas, such as in Toledo (14), Albacete (15), Murcia (16), Almería (17) and Huesca (18).

These data clearly show that the distribution of this phenomenon is much wider than previously claimed, since it is attested not only in the Galician area but also in Leonese and in Andalusian, the latter being the probable source for the transfer of the phenomenon to Latin American Spanish.

Table 1. Data from the ALPI, Cuaderno I: #406. *Estaba medio dormida*

<i>Provinces</i>	<i>Standar medio</i>	<i>Points</i>		<i>No data (ø)</i>	<i>Total</i>
		<i>Vernacular media</i>	<i>Variation medio~a</i>		
1. Coruña	---	14 (100%)	---		14
2. Lugo	6 (42.86%)	8 (57.14%)	---		14
3. Pontevedra	---	13 (100%)	---		13
4. Orense	---	9 (81.82%)	1 (9.09%)	1 (9.09%)	11
5. León	10 (76.92%)	2 (15.38%)	1 (7.69%)		13
6. Zamora	8 (66.67%)	---	4 (33.33%)		12
7. Valladolid	2 (50%)	2 (50%)	---		4
8. Salamanca	8 (72.73%)	2 (18.18%)	1 (9.09%)		11
9. Cáceres	6 (75%)	1 (12.50%)	---	1 (12.50%)	8
10. Huelva	4 (66.67%)	2 (33.33%)	---		6
11. Sevilla	4 (44.44%)	5 (55.56%)	---		9
12. Cádiz	---	4 (100%)	---		4
13. Málaga	7 (87.50%)	1 (12.50%)	---		8
14. Toledo	6 (85.71%)	1 (14.29%)	---		7
15. Albacete	7 (87.50%)	1 (12.50%)	---		8
16. Murcia	7 (70%)	2 (20%)	---	1 (10%)	10
17. Almería	7 (87.50%)	1 (12.50%)	---		8
18. Huesca	12 (92.31%)	1 (7.69%)	---		13
<i>Total</i>	94 (54.34%)	69 (39.88%)	7 (4.05%)	3 (1.74%)	173

As shown in Table 1, the vernacular form with agreeing *media* is either categorical or predominant at 40% of the points taken into consideration (non-Spanish points from the ALPI were excluded, i.e. those from Portugal or from Catalan-Valencian regions). In particular, 82% to 100% nonstandard usage is found in three provinces in the north (1, 3 and 4) and one in the south (province 12). In four provinces (4, 5, 6 and 8) we find variation between standard *medio* and nonstandard *media*. With just 57% nonstandard usage, Lugo (province 2) seems exceptional compared to the rest of Galicia: note that the points in question are adjacent to the region of Asturias, where the nonstandard variant with *media* is not attested.

### 3. Examples of variable *medio* recategorization from text corpora

We also use historical and contemporary data from both Peninsular and Latin American texts (Davies 2001) to analyse some specific cases of adverbial *medio* with (4a) and without (4b) agreement, and to study the semantic interpretations associated with each variant:

- (4) a. *medio tonta* (invariable)  
'rather stupid'.  
b. *media tonta* (agrees fem. sg.)  
'rather stupid'.

In the contrasting examples from (5ab) to (8ab), we see that nearly identical pairs occur both with nonstandard agreement (5a. through 8a) and without (5b. through 8b) in different varieties of written Spanish. Such coexistence of standard and nonstandard forms occurs at times in texts from the same epochs, and even in some cases from the same authors:

- (5) a. como los dioses grandes y medios dioses  
'like the great-M-PL gods and the half-M-PL gods'  
[*Filosofía secreta*, Juan Pérez de Moya (1554)].  
b. donde las ninfas y otros medio dioses se ayuntaban a hacer sus fiestas  
'where the nymphs and other half-M-SG gods gathered to have their parties'  
[*Filosofía secreta*, Juan Pérez de Moya (1554)].  
(6) a. mis medios hermanos son  
'my half-M-PL brothers are > they are my half brothers'  
[*Los cabellos de Absalon*, Pedro Calderón de la Barca (1640)].  
b. sus medio hermanos morían como moscas  
'his half-M-SG brothers died like flies'  
[*La tía Julia y el escribidor*, Mario Vargas Llosa (Peru, 1977)].  
(7) a. era media hermana del ministro de la educación  
'she was half-F-SG sister of the minister of education'  
[*Cuentos*, Guido Rodríguez Alcalá (Paraguay, 1993)].  
b. era hija adúltera de su marido y medio hermana por ende de su desgraciado hijo  
'the illegitimate daughter of her husband, and therefore half-M-SG sister to wretched son'  
[*Cecilia Valdés o La loma del Ángel*, Cirilo Villaverde (Cuba, 1882)].

- (8) a. muchos andaban medio desnudos y sin tener con que se cubrir  
 ‘many went around half-M-SG naked-M-PL and without anything  
 to cover themselves’  
 [Crónica del Perú, Pedro Cieza de León (1551)].
- b. verá sus hijos vagar medios desnudos por estos callejones  
 ‘he will see their children wander half-M-PL naked-M-PL through  
 those alley-ways’  
 [Pedro Sánchez, José María de Pereda (1870)].

The *recategorization* of *medio* appears to be related to the semantic interpretation of the noun or adjective it modifies. Speakers often seem to use adverbial *medio* to introduce a negative quality or characteristic, for example:

- (9) a. no sé si estaría media pesimista  
 ‘I don’t know if I would be half-F-SG (i.e. somewhat) pessimistic’  
 [Habla Culta de Santiago de Chile, M35 (1979/1990)].
- b. la empleada nueva, que tú ves que es media lerda  
 ‘the new maid, who as you see is half-F-SG (i.e. rather) dumb’  
 [Habla Culta de Santiago de Chile, M43 (1979/1990)].

Since it is often used with negative information and pejorative adjectives, adverbial *medio* shows a ‘modal’ meaning, which nuances the negative nature of the adjective by means of the (lesser) degree of ‘somewhat, a bit’. Perhaps out of politeness, when one does not need or want to stress explicitly the negative aspects of an adjective, speakers will often use *medio* to mitigate negative meanings and underscore some undesirable condition. The examples in (10) (taken from oral corpora, Davies 2001) show this tendency clearly:

- (10) a. *media muerta* ‘half dead’.
- b. *media tonta* ‘rather stupid’.
- c. *medios desnudos* ‘semi-naked’.
- d. *medios salvajes* ‘semi-savages’.
- e. *media rara* ‘kind of weird’.
- f. *media estropiada* ‘almost spoiled (of fruit or other foods)’.

It seems that recategorization of *medio* allows us to characterise the object referred to in a [ $\pm$  evaluative] manner, reducing the specificity of a negative term (*tonta*, *rara*, etc.), thus making it seem ‘less strong’. In a formal sense, we can see this sort of ‘quaint agreement’ phenomenon as similar to the agreement of the French adverb *tout* when it modifies adjectives. Furthermore, in colloquial French, adjectives used as adverbs can sometimes agree (Roberge 1999):

- (11) a. Elle sent *bon/ bonne*  
'*She smells good/ good* (fem. sg.)'.
- b. L'avion vole *bas/ basse*  
'*The airplane flies low/ low* (fem. sg.)'.
- c. L'acenseur monte pas ben *haut/ haute*  
'*The elevator does not go very high/ high* (fem. sg.)'.

A similar situation exists in Italian (12ab), and in Spanish too (12c):

- (12) a. María parla *svelto* (m. sg.)  
'*Mary speaks fast*'.
- b. María parla *svelta* (fem. sg.)  
'*Mary speaks briefly*'.
- c. Ella siempre camina *rápida* (fem. sg.)  
'*She always walks fast* (i.e. she is a fast walker)'.

Similarly, there are other Spanish adverbs which have nonstandard 'agreeing' variants, as if they were adjectives, including *puro* 'pure' and *demasiado* 'too much' (Davies 2001):

- (13) a. Pero no me casé por pura perezosa  
'*But I did not get married for pure-F-SG lazy-F-SG (i.e. out of sheer laziness)*'.  
[Habla Culta de Bogotá, M38 (1986)].
- b. de puros cansados y muertos de andar por las nieves e bosques  
'*from pure-M-PL tired-M-PL and dead from walking through snows and forests*'  
[*Cartas del Reino de Chile*, Pedro de Valdivia (1527)].
- (14) a. Mi opinión es que es demasiada avanzada  
'*My opinion is that it is too-F-SG advanced-F-SG*'  
[Habla Culta de La Habana, M11 (1973)].
- b. terminé la carrera demasiada joven  
'*I finished my studies too-F-SG young-SG*'  
[Habla Culta de La Habana, M8 (1973)].

#### 4. Conclusions

Examples of nonstandard agreement presented in this paper (cases where adverbial *media* is used if it were an adjective) show that recategorization of *medio* in Spanish:

i) Does not have a geographic distribution restricted to Latin America or to areas where Spanish is in contact with Galician-Portuguese: it also occurs elsewhere in Peninsular Spanish, especially in part of Andalusia.

ii) Is not a new phenomenon: neither the change of *medio* from an adverb to an adjective nor the variation between invariant and agreeing forms are new developments in modern Spanish: recategorization is a language change phenomenon which is attested since medieval times.

iii) Is not an isolated phenomenon within the grammatical lexicon of Spanish, since there exist analogous examples of agreement with adverbs such as *puro* and *demasiado*.

iv) Is a characteristic feature with alternate forms which can vary not only between speakers, but also within the speech of one individual and within one text.

## References

- Bello, Andrés. 1988[1847]. *Gramática de la lengua castellana destinada al uso de los americanos*. Con las notas de Rufino José Cuervo, ed. Ramón Trujillo. Madrid: Arco/Libros.
- Bosque, Ignacio. 1990. *Las categorías gramaticales. Relaciones y diferencias*. Madrid: Síntesis.
- Bosque, Ignacio. 1999. El nombre común. In *Gramática Descriptiva de la Lengua Española*, dirs. Ignacio Bosque and Violeta Demonte, vol. 1, 3-75. Madrid: Espasa.
- Company, Concepción. 1995. Old forms for new concepts. The recategorization of possessive duplications in Mexican Spanish. In *Historical Linguistics*, ed. Henning Andersen, 77-93. Amsterdam: John Benjamins.
- Cotarelo Valledor, Armando. 1927. El castellano en Galicia. Notas y observaciones. *Boletín de la Real Academia XIV*:82-136.
- Davies, Mark. 2001. *Corpus of Historical Spanish Prose 1200-1900/ Corpus del Español*. Provo: Brigham Young University & National Endowment for the Humanities, <http://www.corpusdelespanol.org>
- Demonte Barreto, Violeta. 2000. Gramática, variación y norma. Una tipología. *Estudios Hispánicos* 17/12:3-49.
- Demonte Barreto, Violeta. 2001. El español estándar (ab)suelto. Algunos ejemplos del léxico y la gramática. In *Actas del II Congreso Internacional de la Lengua Española*. Madrid: Instituto Cervantes, <http://cvc.cervantes.es/obref/congresos/valladolid>
- Dubinsky, Stanley, & Kemp Williams. 1995. Recategorization of prepositions as complementizers. *Linguistic Inquiry* 26:125-137.
- Graell Stanziola, Matilde and Antonio Quilis. 1991. Datos sobre la lengua española en Panamá. In *El Español de América. Actas del III Congreso Internacional de El Español de América*, eds. César Hernández Alonso et al., vol. 2, 997-1005. Salamanca: Junta de Castilla y León.
- Heap, David. 2002. *El Atlas Lingüístico de la Península Ibérica en Internet (ALPI)*. London: University of Western Ontario, [www.alpi.ca](http://www.alpi.ca)



- Kany, Charles. E. 1970[1945]. *Sintaxis hispanoamericana*. Madrid: Gredos.
- Moreno de Alba, José G. 1993. *El español en América*. México: Fondo de Cultura Económica.
- Pato, Enrique. 2005. Selection of informants and its impact on data collected in Spanish dialectology studies. In *International Conference on Methods in Dialectology XII*. Moncton: Université de Moncton (August, 1-5, 2005).
- Pato, Enrique and David Heap. 2004. El proyecto CoRDIALPI (Corpus Retranscrito Digitalmente del Atlas Lingüístico de la Península Ibérica) y el [www.alpi.ca](http://www.alpi.ca). In *VI Congreso de Lingüística General*. Santiago de Compostela: Universidade de Santiago de Compostela (May, 3-7, 2004).
- Quilis, Antonio. 1983. *La concordancia gramatical en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas (CSIC).
- Real Academia Española*. 1972. *Esbozo de una Nueva Gramática de la Lengua Española*. Madrid: Espasa-Calpe.
- Roberge, Yves. 1999. Quaint Agreement and the Theory of Spell-Out. In *Semantic Issues in Romance Syntax*, eds. Esthela Treviño and José Lema, 231-244. Amsterdam: John Benjamins.
- Rojo, Guillermo. 2004. El español de Galicia. In *Historia de la lengua española*, coord. Rafael Cano, 1087-1101. Barcelona: Ariel.