John Benjamins Publishing Company

The generalization of preposition *para* via fusion and ensuing loss of compositionality

Joseph Bauman & Rena Torres Cacoullos
Pennsylvania State University

This study traces two shifts in the distribution of the Spanish preposition *para* 'for, in order to': first, a drop in its allative uses and second, its replacement of the older preposition *por* 'for' with purposive infinitives. These distributional changes of the innovative *para*—across its own contexts of occurrence as well as in its variation with the older *por*—demonstrate the crosslinguistic allative-to-purposive grammaticalization path. Frequent co-occurrence of the source elements, *por* and *a*, foments their coalescence, reflected in changes in the orthographic/phonological form of the fused preposition as it loses structural analyzability. Semantic compositionality, whereby there was a discernable semantic contribution of the allative *a* component, is also lost as early prepositional objects designating persons decline. We find this account of the rise of *para,* based on gradual loss of analyzability and compositionality, to be compatible with the quantitative patterns and more insightful than an opaque and implicitly abrupt notion of reanalysis.

1. Introduction

In (1), from an Old Spanish text (the 14th c. *Zifar*), we observe variation between allative 'to' (1a) and purposive 'for, in order to' (1b) uses of the preposition *para*. In this paper we will see that the proportion of allative uses in [*para* + NP or adverb] occurrences declines on the one hand, and on the other, that the purposive infinitive construction, dominated by *por* for most of the history of Spanish, has become associated with *para*. This pair of changes indicates that the evolution of *para* is a case of the hypothesized cross-linguistic grammaticalization path depicted in (2).

(1) a. *fueron-se*   ***para** la*   *ribera de la mar*
        go.PFV.3PL-REFL to   ART.DEF.F.SG shore of the sea

     b. ***para** se*   *y- r*
        to   REFL go-INF [*Zifar*, 89]
        'they went to the shore to depart'

(2)   Grammaticalization path:
       allative > purposive   (Heine and Kuteva 2002, 39)

We will show quantitative distributions to support our claim that the grammaticalization of *para*, which arises from the fusion of two earlier prepositions—*por* and *a*—involves loss of analyzability and compositionality.

The corpus compiled for this study comprises 17 texts, beginning with *El cantar de mio Cid* (1140–1207) and drawing on two prose texts for each subsequent century up to the 20th. From the 16th century onward, one of these is a peninsular (Spain) text and the other is a chronologically corresponding sample from the *Documentos Lingüísticos de la Nueva España* (New Spain/Mexico). For each century, approximately 1,000 tokens of *por* and *para* with a nominal or infinitive complement were extracted (that is, not extracted were tokens of *por* or *para* followed by the conjunction *que* and a finite verb). Note that counts for *para* include tokens of the form *pora* in the 12th and 13th century. Information on the sampling procedure and exclusions is given in Torres Cacoullos and Bauman (2014, 391–393).

Table 1 shows the texts, token counts, and frequency of *para* relative to *por* by century. Though not spectacular, there does seem to be an increase over time in the relative frequency of *para* on this overall measure. As we will see, superior measures are provided by patterns of co-occurrence with contextual elements, or relative frequencies in linguistic subcontexts.

2.   From allative origins to the decline of spatial uses

The uses of a preposition or the kinds of relations it signals may be classified as spatial, temporal or abstract (e.g., Delbeque 1996, 252). Applying this classification to tokens of *para* with something other than an infinitive complement, usually an NP, we examine here the spatial uses of the preposition. Table 2 shows the verbs modified by *para* in configurations comprised of [VERB + *para* + NP (or adverb)]. The most frequent verbs modified by a *para* phrase in such spatial uses are three verbs of directional motion—*ir* 'go', *venir* 'come', *tornar* 'return'. Furthermore, middle-marked *irse* is somewhat more frequent than unmarked *ir* in the earliest time period (at a ratio of 1.3 to 1, or 56 to 44 tokens). This does not reflect a general fact, since with the preposition *por* the opposite obtains, with *irse* ten times *less* frequent than *ir* (3 to 31 tokens). The preponderance of directional motion verbs, especially *se*-marked motion verbs (Maldonado 1999, 363–373), is consonant with an allative usage encoding movement to or towards a location.

Table 1. Texts by century, token counts (*para* and *por*) and frequency of *para* relative to *por*

Cent	Text*	N	% *para***
12th	*Cid*	370	17%
13th	*Calila*, pp. 91–181	406	17%
	GEI, pp. 5–122	563	
14th	*Zifar*, pp. 9–110	500	24%
	Lucanor, odd numbered *exempla*	572	
15th	*Corbacho*, pp. 67–104 (Primera parte, 1–17), pp. 145–183 (Segunda parte, 1–8)	511	21%
	Celestina, pp. 67–214	556	
16th	*LT*	348	30%
	DLNE 1535–1569, pp. 109–161 (docs. 17–36)	364	
17th	*Quijote II*, odd numbered chapters between 1 and 27	489	23%
	DLNE 1609–1640, pp. 240–347 (docs. 79–129)	495	
18th	*CN/Sí*	339	32%
	DLNE 1790–1810, pp. 611–709 (docs. 258–307)	495	
19th	*Regenta*, Chapters 16, 19, 22, 25	433	33%
	Bandidos, pp. 27–278 (Chapters 1–29)	503	
20th	*Madrid*, pp. 87–290 (transcripts 5–16)	630	32%
	México, pp. 11–172 (transcripts 1–13)	491	

*Editions are listed before the References.
**% *para* in each century combines data from the two texts. 12th- and 13th-century *para* counts include instances of *pora*. Counts of *pora—para* are, respectively, 63–1 in the *Cid*, 78–1 in *GEI*, 1–83 in *Calila*.

Table 2. Most frequent motion verbs modified by *para* and their proportion of the preposition's spatial uses

Century	*Ir*	*Irse*	*tornar(se)*	*venir(se)*	%
12th–14th	44	56	13	19	80% (132/164)
15th–16th	6	0	0	1	29% (7/24)
17th–20th	7	8	0	4	40% (19/47)

The grouping of the centuries in Table 2 into three time periods, namely 12th–14th, 15th–16th and 17th–20th, emerged once we considered each century separately. In the 12th–14th-century data, the three verbs of directional motion together constitute 80%

(132/164)—*ir(se)* alone close to two-thirds—of all spatial instances of *para*. Repeated co-occurrence with *ir(se)* is important for the evolution of *para*, in light of the proposal from typological studies that "in the presence of 'go'-verbs, allatives frequently take on a purposive reading" (Rice and Kabata 2007, 459) An example would be German *zu,* whose "allative use [ ] evolves into a purposive one" (Lehmann 2002, 6), as in (3) (intermediate stages not shown).

(3)   German *zu*: allative > purposive > subordinator of infinitive

(adapted from Lehmann 2002, 6)

a.   *Der Prinz begab sich **zur** Königin*
'The prince betook himself to the queen'

b.   *Der Prinz begab sich **zum** Jagen*
'The prince betook himself to hunting'

c.   *Der Prinz entschied sich **zu** jagen*
'The prince decided to hunt'

Formulations such as "take on a purposive reading" or "evolves into a purposive" do not mean an abrupt or even linear replacement of one use by another. We stress that from the earliest texts, allative and purposive uses coexist, as illustrated in (1) above. Rather, "semantic change […] should be manifested in changing distribution and co-occurrence patterns" (Torres Cacoullos and Schwenter 2005, 357).

One measure is the distribution of *para* across its contexts of occurrence. We find that, while the proportion of temporal uses (e.g., *para mañana* 'for, by tomorrow') has remained steady at approximately 5% to 10% throughout the centuries, there is a notable shift in the proportion of spatial uses. The line marked with diamonds in Figure 1 shows the proportion of spatial uses of *para* by century (not counting occurrences with an infinitive complement). Even in the earliest texts *para* expresses spatial relations in (approximately) half (43%–52%) of its occurrences with an NP (or adverbial) object, that is, we do not find an initial period in which *para,* or its precursor *pora*, exclusively or even mostly had a spatial sense. Nevertheless, after the 14th century the proportion of spatial uses is no greater than 10%.

Comparison with *por* confirms that the decline of spatial uses is not a mere accident of genre or topic (for example, we would expect higher proportions of spatial uses in epics). The line marked with squares in Figure 1 shows the corresponding distribution of *por* tokens. In contradistinction to the decline of spatial uses with *para*, the stability of *por* in spatial uses is evident, with a proportion of approximately one-fifth (ranging from 10% to 30%).

In summary, the allative origins of *para* are indicated by co-occurrence with directional motion verbs, often *se* (middle)-marked, with spatial uses constituting approximately half of the preposition's tokens with a NP or adverbial complement.
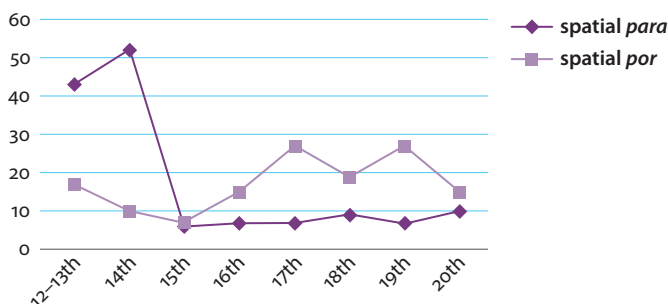
Figure 1. Proportion of spatial uses of *para*, compared with spatial uses of *por* (Ns *para*: 12–13th 75/173, 14th 89/170, 15th 7/120, 16th 8/115, 17th 9/122, 18th 12/134, 19th 9/122, 20th 26/261; Ns *por*: 12-13th 164/959, 14th 71/724, 15th 51/713, 16th 55/370, 17th 160/587, 18th 92/487, 19th 157/575, 20th 85/579.)

After the 14th century, *para* with a nominal complement appears mostly with non-spatial, abstract senses.[1]

3.   Generalization in purposive infinitive constructions

A second major quantitative change in co-occurrence patterns concerns infinitive complements. Here it is not distribution across contexts of occurrence that provides an enlightening measure, but variation with respect to the older preposition, *por*.
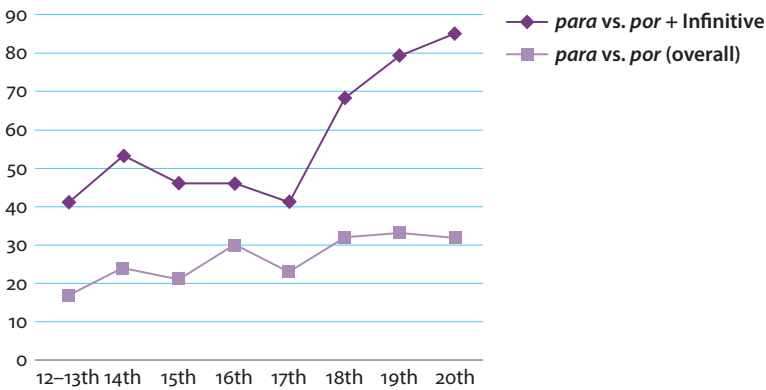
Table 3 depicts the frequency of *para* relative to *por* with an infinitive complement. The two prepositions are fairly evenly distributed (~50%) in this context until the 17th century, after which the rate of *para* increases, such that it becomes two to five times greater than that of *por* (68%, 79% and 85% in the 18th, 19th and 20th centuries, respectively).

Figure 2 compares the increasing frequency of *para* relative to *por* overall (seen earlier in Table 1), in the line marked with squares, with that in the particular context of infinitives (Table 3), in the line marked with diamonds. It is clear that the frequency increase of the newer preposition has occurred disproportionally precisely in [+ infinitive] constructions.

---

1.   The 14th century is when *para* definitively displaces *pora* (Riiho 1979, 232)

Table 3. Frequency of [*para* + infinitive] relative to [*por* + infinitive]

Century	% *para*: relative to *por* with an infinitive	N
12th–13th	41%	77/187
14th	53%	84/159
15th	46%	100/219
16th	46%	92/201
17th	41%	111/272
18th	**68%**	127/187
19th	**79%**	183/232
20th	**85%**	200/235

Figure 2. Increasing relative frequency of *para* vs. *por* overall (Table 1) and more striking increase in relative frequency of [*para* + infinitive] vs. [*por* + infinitive] (Table 3)

But is the displacement of *por* by *para* with infinitives a real linguistic change, or do the rate shifts merely reflect shifts in what is talked about or how it is talked about, that is, change in cultural context? (For such a scenario, see Myhill (1995) on American English modals.) In present-day Spanish, infinitive constructions with *por* mean something different from those with *para*, the former generally expressing cause (or reason) and the latter, purpose. For example, in (4a), with *por*, the subject felt guilty *because* he married off someone, whereas in (4b), with *para*, the subject needed money *in order to* retrieve his clothing. It is not inconceivable, then, that the increasing rate of *para* with infinitives reflects a shift toward more talk of purposes than of causes.

(4)   a.   [verb + **por** + infinitive] = cause (reason)
           *Se creía […] culpable **por** haber casado a Tules* [19th c., *Bandidos*, 1.159]
           'He thought himself […] guilty **for** having married off' Tules'

b.   [verb + ***para*** + infinitive] = purpose
*necesitaba diez pesos **para** <u>sacar</u> su ropa empeñada* [19th c., *Bandidos*, 1.154]
'he needed money **in order to** <u>retrieve</u> his clothing that had been pawned'

To probe this issue, we look closely at cases where the prepositional phrase modifies a verb, i.e. at [verb + *por/para* + infinitive], since this is the locus of the generalization of [*para* + infinitive].[2] As a replicable measure, the purpose sense is operationalized in terms of the temporal reference of the situation: counted as having a purpose sense were cases in which the situation referred to by the infinitive is posterior to that of the main (finite) verb, as in (5). Figure 3 shows the distribution of *por* and *para* in *purposive* infinitive constructions in a sample of the texts. We see that in the 15th-century *Celestina* and 17th-century *Quijote*, purposive infinitive complements are evenly distributed between the two prepositions (50% (35/70) and 46% (28/61) for *para* in the two texts, respectively). That is, in Old and Golden Age Spanish, infinitives with *por* could, and robustly did, express purpose (as in (5)). In the 19th-century *Regenta* (Spain) and *Bandidos* (Mexico), however, the relative frequency of *para* in this same context is up to 90% (111/123). Thus, *para* has generalized in the purposive infinitive construction, largely replacing the older preposition, in support of genuine linguistic change.

(5)   [verb + ***por*** + infinitive] = purpose
*También quiere a mí engañar como a mi amo **por** <u>ser</u> rica* [15th c., *Celestina*, 5.173]
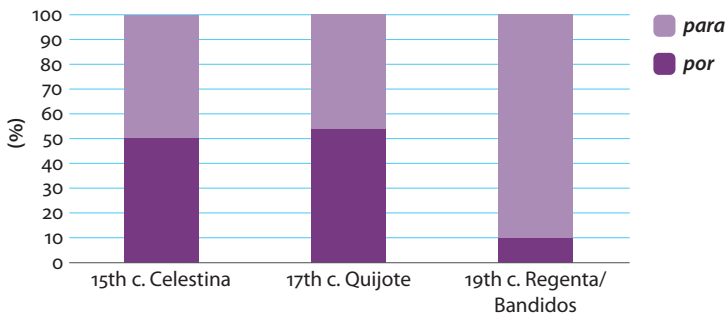'She wants to trick me as well as my master **in order to** <u>become</u> rich'

Figure 3. Rate of *para* relative to *por* with a purposive infinitive complement ([*para* + infinitive$_{\text{purposive}}$] vs. [*por* + infinitive$_{\text{purposive}}$])

---

2.   The proportion of tokens of [*para* + infinitive] modifying a nominal element is steady over time at approximately one-fourth, whereas for *por* it has remained at 5% or lower. Nouns recurring with a [*para* + infinitive] complement are *esfuerzo(s), facilidad, fuerza(s), licencia, motivo(s), tiempo, valor* and adjectives *aparejado, bueno, eficaz, hábil, mejor, necesario, suficiente, útil* , as in *sin <u>fuerzas</u> **para llamarle*** 'without strength to call him' [*Regenta*, 2.321].

Still, it might be argued that even in purposives there may be a meaning difference, such that purposive infinitive clauses with *por* express a nuance of "underlying motive or incentive" that is absent from *para* (Bolinger 1945, 20).[3] How can we confront such a claim? The working hypothesis of the variationist method is that "within a given locus of variability, or *variable context*, [....the] competing variants will occur at greater or lesser rates depending on the features that constitute the context" (Poplack 2001, 405). If linguistic forms mean something different, they should be preferred in different (sub)contexts. In other words, they should occur at higher or lower than average rates in certain (sub)contexts (see also Aaron and Torres Cacoullos 2005, 615).

Therefore, we seek corroboration of the generalization of *para* in purposive infinitive constructions to the detriment of *por* indicated above in Figure 3 by comparing the rates of the prepositions in two particular subcontexts that are compatible with purposive meaning. One is [motion verb + *por/para* + infinitive], i.e. where *por* or *para* with an infinitive complement modifies *acercarse* 'to approach', *andar* 'to walk', *correr* 'to run', *descender* 'to descend', *dirigirse* 'to set off, go', *entrar* 'to enter', *huir* 'to flee', *ir* 'to go', *llegar* 'to arrive', *moverse* 'to move', *salir* 'to leave', *seguir* 'to follow', *venir* 'to come', *volver* 'to return' and other intransitive motion verbs, as in (6). These verbs favor use of [*por* + infinitive], with *para* at 32% (27/85) until the 17th century, after which we observe a reversal, with *para* up to 90% (55/61).

(6)   variation in [motion verb + *por/para* + infinitive]

a.   *quiero yr ala cort, **por de mandar** myos derechos*
'I want to go to the court **in order to** demand my rights.' [*Cid*, 3079]

b.   *dixo a su hermano que ella quería yr con él aquella noche **para traer aquello***
'she said to her brother that she wanted to go with him that night **in order to** recover that.' [Lucanor, Ej. 47]

A second subcontext providing evidence for genuine linguistic change is that in which the subject NP has a human referent. If there is a meaning difference between [*por/para* + infinitive] such that *por* expresses a nuance of "underlying motive or incentive" (as claimed by Bolinger 1945, 20), we would expect human subjects to favor the use of *por* and inanimate subjects to favor *para*, since inanimate subjects are incapable of having motives or intentions. For example, with the inanimate subject 'fortune' in (8a), the main verb is a stative (*es favorable* 'is favorable') and the usage is that of purpose in

---

3.   We take here the stance that the meaning(s) associated with a form are evident in its usage (i.e. usage and meaning are directly linked in usage-based theory (Bybee 2010). That is to say that the contexts in which a form appears, given frequent and sustained application to a given usage, can effect a gradual but permanent change in the meaning(s) that are assumed to be inherent to that form.

the sense of the utility of a thing. With human subjects (7), however, *para* would seem to express the purpose of an action with a sense of intention. Note that there would not appear to be a justification for considering the instance with *por* with a volitional human subject in (7a) as conveying more of a sense of underlying motive than the instance with *para* in (7b) (besides the circular argument that *por* itself has such a meaning).

(7)   [human subject + verb + **por/para**+ infinitive$_{purposive}$]

   a.   *quiso turbarme* **por** *oirme decir otras docientas patochadas*
        [*Quijote* II, 7.680]
        '<u>he</u> tried to upset me **so as to** <u>hear</u> me say another two hundred follies'

   b.   *quería vencerla,* **para** *no padecer tanto* [*Regenta*, 2.22]
        '<u>he</u> wanted to overcome it [his ambition], **so as to** <u>avoid suffering</u>
        so much'

(8)   [inanimate subject + verb + **por/para**+ infinitive$_{purposive}$]

   a.   *para qué es la* <u>*fortuna*</u> *favorable y próspera sino* **para** <u>*servir*</u> *a la honrra*
        [*Celestina*, 2.130]
        'for what is <u>fortune</u> favorable and propitious if not **to** <u>serve</u> honor'

   b.   *toda la* <u>*natura*</u> *se remiró* **por** *la* <u>*hazer*</u> *perfecta* [*Celestina*, 6.191]
        'all of <u>nature</u> exerted itself **to** <u>make</u> her [Melibea] perfect'

Figure 4 shows the distribution of *por* and *para* in purposive infinitive constructions with inanimate vs. human subjects (of the main verb). It does appear that inanimate subjects have always favored the choice of [*para* + infinitive] (top figure) (though Ns are low). However, while human subjects favor [*por* + infinitive] in the 15th-c. *Celestina* and 17th-c. *Quijote*, with *para* at 44% (48/110), the rate of *para* is 89% (102/114) in the 19th-c. *Regenta* and *Bandidos* (bottom figure). This reversal is a second measure of linguistic change: whereas in earlier times intentions of human subjects were expressed with *por,* now this function is served by *para*.

   In summary, we first observed a reversal in the relative frequency of *por* and *para* in infinitive constructions overall (Figure 2). We then observed a parallel reversal more particularly with a *purposive* infinitive complement (counting as purposives those that are temporally posterior to the main verb) (Figure 3). Finally, we considered two replicable measures—the rate of *para* in infinitive constructions with motion main verbs and with human subjects—which again show a reversal of relative frequencies of *por* and *para* (Figure 4). The conclusion is that *para* has generalized as a purposive infinitive marker, as we may represent in (9).

(9)   [VERB + *para* + infinitive]$_{purpose}$

Together with the decline of allative uses (Section 2), this is solid evidence for the postulated allative > purposive evolutionary path (exemplified from other languages in Heine and Kuteva 2002, 39–40). Such changes are often thought of as reanalysis (e.g.,
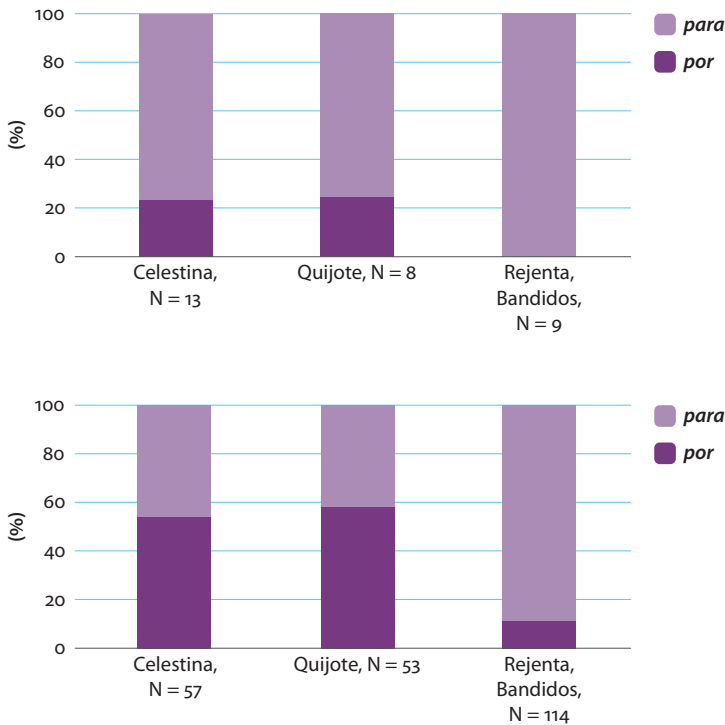
Figure 4. Rate of *para* relative to *por* with a purposive infinitive complement modifying a verb with an inanimate subject (top) or a human subject (bottom)

Campbell 1998, 284), which has been conceived as abrupt in language change. In the following section, we argue for an alternative view based on chunking and ensuing loss of analyzability and compositionality.

4.   Fusion and loss of internal analyzability: *Por + a > para*

It is widely hypothesized that *para* arises from fusion, or coalescence (Haspelmath 2011), of *por* (or its Latin antecedents PER and PRO (Riiho 1979, 13-28)) and *a(d)*. Evidence comes from 12th–13th-century examples in which *por* and *a* appear together in non-agglutinated form.[4] Example (10) illustrates variation between the non-agglutinated and the agglutinated form in the 13th-century *General Estoria, Primera parte* (*GEI*), in a near-identical context.

---

4.   30 tokens of *por + a* separated by a space (i.e. not agglutinated in the 600,000-word digital version of the *GEI* (Kasten, Nitti and Jonxis-Henkemans 1997).

(10)   *por* + *al* vs. *poral* variation
  *diz q⟨ue⟩ agun bien es **por al** om⟨n⟩e en no⟨n⟩ seer pecador. & maguer que*
  *non es sa⟨n⟩cto. como diz q⟨ue⟩ es otrossi algo **poral** om⟨n⟩e del qui non puede*
  seer Rey. [13th c., *GEI*, fol. 57v]
  'It is said that there is some good **for** man in not being a sinner even though
  he is not holy, as it is said that this is furthermore something **for** the man
  who cannot be king.'

The majority of the tokens of *por* + *a* separated by a space in the *GEI* occur with a noun as the object of the preposition and with the definite article fused with *a* in *al*, as in the first line in (10) above. That the construction which most resisted the fusion of *por* and *a* is with the masculine definite article *el* is perhaps due to countervailing force of the contraction of *el* to the preposition *a* (the contracted form *poral* persists into 17th-century texts (Riiho 1979, 236)).

  Nevertheless, we also find examples of non-agglutinated *por* + *a* in other contexts, for example, with a pronoun, as in ***por ami*** 'for me' (12th c. *La Fazienda de Ultra-mar* 13vA25, Ex.6, 7) (Dave McDougall, p.c.) or preceding an infinitive, as in ***por a yr a Egipto*** 'to go to Egypt' (13th c. *GEI*,fol. 108v). Thus, although tokens with a space between *por* and *a* are already a tiny minority, there are enough to suggest that scribes utilized the non-agglutinated combination with some regularity in 13th-century texts.

  Indeed, in 13th-century Spanish texts, there is a range of antecedent forms for *para*. Most prominent is the agglutinated form *pora*, as illustrated in (11) and (12). The relative frequency of *pora* with respect to *para* is reported to decrease precipitously from 83% to 15% from the first to the second half of the 13th century (Riiho 1979, 232). Other candidate antecedent forms are *pera* and *perad* (García de Diego 1951, 128).

(11)   Earlier form *pora*…
  a. *Vansse **pora** San Pero* [12th c., *Cid*, v. 294]
   'They go **to** San Pedro'
  b. *estas serien despues **pora** comer **pora** ell omne* [13th c., *GEI*, fol. 12v]
   'these would be then **for** man **to** eat'

(12)   …in variation with *para*
  a. ***para** Calatayuch quanto puede se va* [*Cid*, v. 774–775]
   '**to** Calatayud as soon as he can he goes'
  b. *fuel aparta⟨n⟩do toda uia **para** si & alos suyos.* [13th c., *GEI*, fol. 5r] [5]
   'he divided it [the land] **for** himself and (to) his own'

However, we find inconsistency across and within editions of Old Spanish texts in the realization of '*para*' and its variant forms. For example, in reproductions of the

---

5. In …***para** si & alos suyos* [*GEI*, fol. 5r] (Example 12), *para* and *a* appear to compete in the same context.

original manuscripts of the 13th-century *GEI* there already appear unequivocal tokens of *para* (13).

(13)   *para* in 13th-century original manuscript
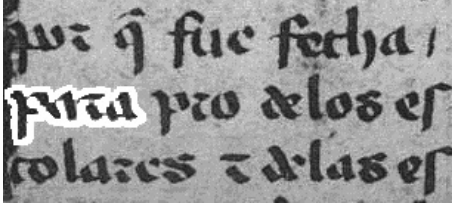
…*por q̄ fue fecha* **para** *pro delos escolares*…   [Solalinde 1930, 287]
'by which it was made for the good of the scholars'

In the same text, scribes also used the form depicted in (14). Here the preposition appears as an ambiguous abbreviation without the interior graphemes, displaying only the initial *p* and the final *a* (rather than representing loss of the second syllable (see Company 1994, 11)).
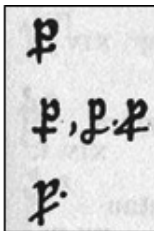
(14)   'para' in same 13th-century original manuscript
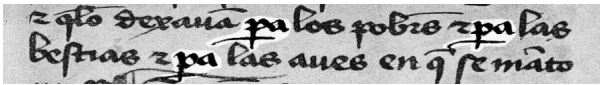
et q̄lo dexauā **pa** *los pobrē et* **pa** *las bestias et* **pa** *las aues*…
   [Solalinde 1930, 424]
'…and that he left it for the poor and for the beasts and for the birds'

The abbreviation with a horizontal bar either crossing or connecting with the descender (the vertical line) of the *p*, depicted in (14) was a convention in the transcription of Latin *per* and *pro* (Cappelli 1899/1990, 257). In Cappelli's dictionary of abbreviations, the horizontal bar does not cross the descender in the abbreviation for *pro,* but merely connects with its left side. Thus, the abbreviations used for *per* more closely resemble the form of the *p* that is encountered in 13th-c. Spanish manuscripts, as in (14) above.

(15)   Abbreviations for *per:*   Abbreviations for *pro:*
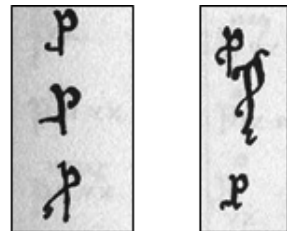
We find another example of this abbreviation, this time involving the adjective *perdonadas* (the adjectival form of the verb *perdonar* 'to pardon', which is unequivocally composed of the prefix *per-* and the root *donar* 'to give'), shown in (16). Nevertheless,

this unambiguous abbreviation of *per-* using a *p* with a crossbar does not mean we must conclude that the examples of abbreviated *pa* necessarily represent *pera, pora,* or some other specific variant. Rather, it appears that this abbreviation was used to represent some vowel-/r/ combination, including the combinations of a vowel and /r/ in *pora, para,* and *per.*[6]

(16)
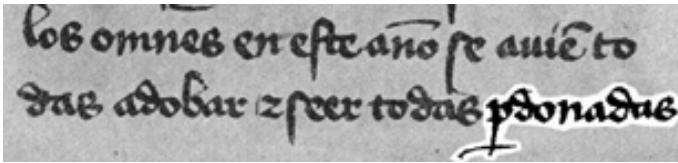
*los omnes en este año se auiē todas adobar et seer todas **pdonadas***

[Solalinde 1930, 424]

'the men in that year all had to dress and be pardoned'

It is clear, then, despite the vexing problem of manuscript editions and graphic conventions, that for a period of time non-agglutinated (*por a*) and agglutinated (*pora* but also *para*) orthographic variants were used contemporaneously. This orthographic evidence indicates that the combination of the prepositions *por* and *a* as independent elements may have persisted into early (13th-century) Spanish. Rather than a linear, consecutive evolution (*per/pro + ad > pora > para*), the coexistence of these forms in the same text indicates that the process of the fusion of the two (groups of) structural elements was not only gradual but also characterized by variation (cf. Weinreich, Labov and Herzog 1968).

The phonological course by which the first vowel in forms spelled *pera* and *pora* ultimately became /a/ in the modern *para* has been addressed by several scholars, appealing to various phonological processes, include lowering of the *e* in *pera* before a rhotic (e.g., García de Diego 1951, 128) and vocalic assimilation (e.g., Hanssen 1945, §726). Regardless of the exact course of events, the ultimate resolution of the variation between *pora* and *para* in the single modern form *para* is further evidence of the formal fusion of the erstwhile *per/pro* and *ad.* It is telling that as *para* wins out over *pora* by the end of the 14th century, the agglutinated form is already effectively unrecognizable (from both a phonological and an orthographic perspective) as consisting of two independent elements.

A requirement for the creation of a new unit is frequent co-occurrence of its erstwhile component parts.[7] From the perspective of a usage-based approach to grammatical forms, a mechanism for the creation of constituent structure is the chunking of

---

6.   Or even a consonant-vowel combination, in light of the frequent transposition of the *r* in Old Spanish (Corominas 1980–1991: see *por*).

7.   It could be argued that an additional requirement is the "semantic coherence" of the elements constituting the combination (Bybee 2010, 138). There has been discussion of the antecedent(s) of *por* (Latin *per* and/or *pro*), and their respective semantic contributions in

a sequence of morphemes that results from frequent repetition of the sequence (Bybee 2010, 34). In other words, "Items that are used together fuse together" per Bybee's Linear Fusion Hypothesis (Bybee 2002, 112). The hypothesis of chunking predicts that for the sequence *por + a* to result in a fused unit—whereas *por* in combination with other prepositions does not—the co-occurrence of *por + a* must be more frequent than other combinations.
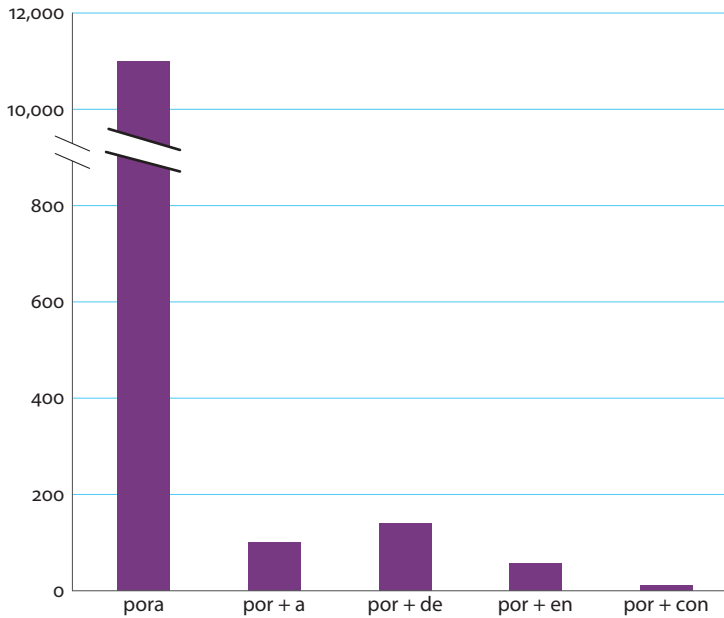

Figure 5. Token frequency of the sequence *por* + preposition (*Corpus del español* counts for 1200s)

Figure 5 shows the token frequency of *por* followed by another preposition for 13th-century texts drawn from the online corpus of Davies (2002-). For the purposes of tabulating co-occurrence of *por + a* vs. *por* + another preposition we count together *pora* ((already) agglutinated) and *por + a* (not-yet agglutinated), given the variation in this period between the agglutinated and non-agglutinated forms (example (10), above). When combined with the already orthographically fused *pora*—more than 10,000 tokens, compared with fewer than 200 cases of *por + de*—it is clear that the *por + a* sequence is indeed of high frequency compared to other sequences, as predicted

---

combination with *ad* (see Torres Cacoullos and Bauman 2014, §13.3 for a recent review). For an assessment of the semantic compositionality of *para*, see Section 5 below.

by the usage-based hypothesis of chunking with repetition. That this fusion was likely complete by the end of the 14th century is suggested by the disappearance of the *pora* variant. This leaves only *para,* a form no longer analyzable as a complex structure.

5.   Compositionality of *para*, and its loss

We saw (Section 2 above) that one change in the distribution of *para* was the quantitative decline of spatial uses after the 14th century. When we look more closely at those early spatial uses, we see that the change further involves the loss of a particular kind of spatial use. In the earliest texts we find examples in which the subject of the motion verb actually arrives at or reaches their destination. For example, in (17) below, *para* indicates the terminal point of the movement, a use that we associate with the preposition *a* in present-day Spanish. In (17a) the subject not only goes 'toward' the *posada* but actually reaches it, as verified by the fact that he speaks to the person there. In (17b) the subject came not 'toward', but 'to', Toledo.

(17)   [motion verb + *para* + NP] = to a location
   a.   *E el pleteo con ellos e fuese **para** la posada e dixole su muger commo auia pleteado con los marineros* [14th c., *Zifar*, 87]
          'And he argued with them and went **to** the inn and told his wife how he had argued with the seamen'
   b.   *et por ende vínose **para** Toledo para aprender de aquella sciençia* [*Lucanor*, 94]
          'And therefore he came **to** Toledo to learn that science'

The endpoint of the motion may even be a person. Cases in which the object of *para* has a human referent are exemplified in (18). Again, in present-day Spanish we would not expect *para* in this context (but *a*).

(18)   [motion verb + *para* + NP$_{\text{specific human}}$]
   a.   *Venimos nós **para** ti que nos consejes* [13th c. *Calila*, 144]
          'We came **to** you so that you may advise us'
   b.   *E el moço se fue **para** su padre, e dixo la respuesta* [14th c. *Zifar*, 21]
          'And the boy went **to** his father and told him the answer'
   c.   *y assi me fuy **para** mi amo, que esperandome estaua.* [16th c. *LT*, 7]
          'and thus I went **to** my master, who was waiting for me'

Figure 6 shows the proportions of spatial uses of [*para* + NP] with prepositional objects that are persons, as in (18) above, as opposed to places (as in (17)). Object NPs with a human referent as the endpoint of motion constitute a full third, 35% (50/144), of spatial instances of *para* in 13th–14th century texts.
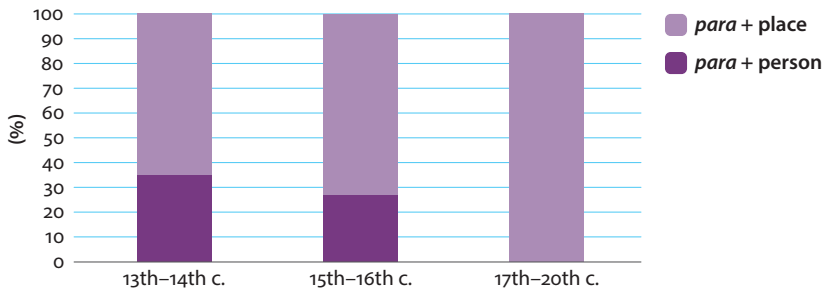
Figure 6. Spatial uses of [*para* + NP]: proportion of object NPs designating persons vs. places

In this motion-verb *para* construction with a destination that is a person, the human referent of the object appears as a personal pronoun (N = 20) or definite full NP (in which the determiner is a definite article (N = 16), possessive (N = 13), or demonstrative).[8] Furthermore, this NP with a human referent overwhelmingly appears in singular number (92%, 46/50, of 13th–14th c. tokens). From these nominal features we may infer that the referent is specific and individuated. This is precisely the kind of referent we would expect the preposition *a* to co-occur with, in accordance with its use as a dative marker (indirect objects tend to be human) and progressively also as an accusative marker for direct objects referring to persons (on the latter, see Company Company 2002).

We submit that this early *para* use with prepositional complements designating persons indicates that the semantic contribution of the preposition *a* was still discernible in the new preposition, that is, it constitutes evidence for a degree of compositionality in the beginnings of the new preposition (cf. Riiho 1979, 99). Following Bybee (2010, 44–45; see also Croft and Cruse 2004, 250–253; Langacker 1987, 292), compositionality is a semantic measure having to do with transparency of meaning, and refers to the degree to which the meaning of the whole is predictable from the meaning of the component parts (e.g. *hopeful* is more compositional than *awful*).[9] Figure 6 indicates that cases are still found in 15th–16th texts (4/15). But the [motion verb + *para* + NP$_{specific human}$] construction disappears from the corpus after the 16th c. (0/51). The disappearance of *para* object NPs with a human referent as the endpoint of motion may be taken as evidence that *a* no longer makes an independent semantic

---

8.   In *Calila*, we count as human the personified animals.

9.   In contrast with the semantic criterion of compositionality, *analyzability* is a morphosyntactic parameter, referring to the degree to which the internal structure and individual parts are recognizable; for example, while *pull strings* is not fully compositional because it has a metaphorical meaning, it is analyzable in that speakers are assumed to recognize an individual verb and its noun complement (Bybee 2010, 45).

contribution, thus demonstrating loss of compositionality, as *a* is completely absorbed into the new, chunked, preposition.

In summary, in addition to the presence of non-agglutinated *por* + *a* orthographic variants, another piece of evidence that *para* is the outcome of the fusion of two separate prepositions, one of which was *a*, comes from early compositionality of semantic content. We have shown how the semantic measure of compositionality can be operationalized in terms of the contexts of use of a linguistic form. In particular, we considered the kinds of object NPs with which *para* co-occurs in spatial relations. A robust [motion verb + *para* + $NP_{\text{specific human}}$] construction in early texts indicates early compositionality. The decline and disappearance of this construction indicates erosion of the semantic contribution of *a* as it is absorbed into the new preposition, and thus loss of compositionality of the whole.

6.   Conclusion: Loss of analyzability and compositionality in grammaticalization

In tracking its generalization, we have seen two principal distribution shifts in the history of *para*, which is a newcomer among Spanish prepositions. First is the general decline of allative uses that is observed following the 14th century, after which the proportion of spatial uses of any kind fails to rise above one tenth of the occurrences of the preposition with a nominal (or adverb) complement. The second change, beginning in the 18th century, concerns infinitive complements, with which *para* replaces *por*, to become the majority variant for the expression of purpose. This pair of distributional changes provides a quantitative argument for allative-to-purposive evolution, a change that has been put forward as a cross-linguistic grammaticalization path (Heine and Kuteva 2002). Thus, change is observable both in the distribution of the newer preposition across its contexts of occurrence and in patterns of variation with respect to the older preposition.

The generalization of *para* as a purposive proceeds from the fusion of its erstwhile component parts. Support for the origin of *para* in the fusion of the sequence of *por* (< Latin *per* and/or *pro*) and *a* (< *ad*) comes from early analyzability and compositionality. Structural analyzability is indicated by instances, though rare, of non-agglutinated *por* + *a* in 13th-c. manuscripts. Semantic compositionality, which involves the independent semantic contribution of the preposition *a*, is discernable in an allative construction where the prepositional object designates a destination that is a person (rather than a location).

The new preposition is consolidated via subsequent loss of analyzability and compositionality. On the one hand, coalescence of the two prepositions is correlated with their frequent co-occurrence and is manifested in sound change in the new fused

form. We take this as evidence for loss of analyzability. On the other hand, loss of compositionality is inferred from the decline and eventual disappearance of *para* objects designating a person.

In summary, allative > purposive grammaticalization of *para* proceeds via loss of morpho-syntactic analyzability, indicated by graphemic (and sound) change, and loss of semantic compositionality, operationalizable through co-occurrence patterns. In the loss of analyzability and compositionality the evolution of *para* is very similar to that of present-day complex prepositions such as Spanish *a pesar de* (Torres Cacoullos and Schwenter 2005; Torres Cacoullos 2006) and English in *spite of* (Beckner and Bybee 2009; Bybee 2010, 136–147). The conclusion is that what is often referred to as 'reanalysis' and conceived of as an abstract and abrupt change is the outcome of gradual processes of loss of analyzability and compositionality.

Corpus of texts (in chronological order, except for DLNE)

[*Cid*] Anonymous, *Cantar de mio Cid. Texto, gramática y vocabulario*, volume 3: *Texto*, paleographic edition of Ramón Menéndez Pidal. Madrid: Espasa Calpe, 1944–1945.

[*Calila*] Anonymous, *Calila e Dimna*, edition of Juan Manuel Cacho Blecua and María Jesús Lacarra. Madrid: Castalia, 1984.

[*GEI*] Alfonso X, *General estoria. Primera parte*, edition of Lloyd Kasten, John Nitti and Wilhelmina Jonxis-Henkemans, *The Electronic Texts and Concordances of the Prose Works of Alfonso X, El Sabio*. Madison: Hispanic Seminary of Medieval Studies, 1997.

[*Zifar*] Anonymous, El libro del cavallero Zifar, edition of Charles Ph. Wagner. Ann Arbor: University of Michigan Press, 1929.

[*Lucanor*] Don Juan Manuel, *El conde Lucanor o Libro de los enxiemplos del conde Lucanor et de Patronio*, edition of José Manuel Blecua. Madrid: Castalia, 1969/1971.

[*Corbacho*] Alfonso Martínez de Toledo, *Arcipreste de Talavera o Corbacho*, edition of Michael Gerli. Madrid: Cátedra, 1979/1992.

[*Celestina*] Fernando de Rojas, *La Celestina*, edition of Dorothy S. Severin. Madrid: Cátedra, 1993.

[*LT*] Anónimo, *Tri-linear edition of* Lazarillo de Tormes *of 1554 (Burgos, Alcalá de Henares, Amberes)*, edition of Joseph V. Ricapito. Madison: The Hispanic Seminary of Medieval Studies, 1987.

[*Quijote*] Miguel de Cervantes, *Don Quijote de la Mancha*, in *Obras completas*, volume 2, edition of Francisco Sevilla Arroyo and Antonio Rey Hazas. Madrid: Alianza Editorial-Centro de Estudios Cervantinos, 1996.

[CN/*Sí*] Leandro Fernández de Moratín, *La comedia nueva. El sí de las niñas*, edition of John Dowling and René Andioc. Madrid: Castalia, 1968.

[*Regenta*] Leopoldo Alas «Clarín», *La Regenta*, edition of Gonzalo Sobejano. Madrid: Castalia, 1981/1982.

[*Bandidos*] Manuel Payno, *Los bandidos de Río Frío*, in *Obras completas*, edition of Manuel Sol. México: Consejo Nacional para la Cultura y las Artes, 2000.

[*Madrid*] Manuel Esgueva y Margarita Cantarero (eds.), *El habla de la ciudad de Madrid: materiales para su estudio*. Madrid: Consejo Superior de Investigaciones Científicas, 1981.

[*México*] Juan M. Lope Blanch (coord.), *El habla de la ciudad de México. Materiales para su estudio*. México: Universidad Nacional Autónoma de México, 1971.

[*DLNE*] Concepción Company Company, *Documentos lingüísticos de la Nueva España. Altiplano central*. México: Universidad Nacional Autónoma de México, 1994.

References

Aaron, Jessi Elana, and Rena Torres Cacoullos. 2005. "Quantitative Measures of Subjectification: A Variationist Study of Spanish *salir(se)*." *Cognitive Linguistics* 16(4): 607–633. DOI: 10.1515/cogl.2005.16.4.607

Beckner, Clay, and Joan Bybee. 2009. "A Usage-based Account of Constituency and Reanalysis." *Language Learning* 59: 27–46. DOI: 10.1111/j.1467-9922.2009.00534.x

Bolinger, Dwight L. 1945. "Purpose with Por and Para." *Modern Language Journal* 28(1): 15–22. DOI: 10.2307/317177

Bybee, Joan L. 2002. "Sequentiality as the Basis of Constituent Structure." In *The Evolution of Language out of Pre-language*, ed. by T. Givón, and Bertram F. Malle, 109–134. Amsterdam: Benjamins. DOI: 10.1075/tsl.53.07byb

Bybee, Joan L. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511750526

Campbell, Lyle 1998. *Historical Linguistics: An Introduction*. Cambridge, MA: MIT Press.

Cappelli, Adriano. 1889/1990. *Lexicon Abbreviaturarum: Dizionario di abbreviature latine ed italiane*. Sexta edición. Trento: Ulrico Hoepli Editore.

Company Company, Concepción. 2002. "Grammaticalization and Category Weakness." In *New Reflections on Grammaticalization*, ed. by Ilse Wischer and Gabriele Diewald, 201–215. Amsterdam: John Benjamins. DOI: 10.1075/tsl.49.14com

Company Company, Concepción. 1994. "Introducción", in *Documentos lingüísticos de la Nueva España*, Mexico: Universidad Nacional Autónoma de México, 1–19.

Corominas, Joan. 1980–1991. *Diccionario crítico-etimológico castellano e hispánico*, con la colaboración de José Antonio Pascual. Madrid: Gredos.

Croft, William and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511803864

Davies, Mark. 2002-. *Corpus del Español: 100 million words, 1200s-1900s*. Available online at http://www.corpusdelespanol.org.

García de Diego, Vicente. 1951. *Gramática histórica española*. Madrid: Gredos.

Hanssen, Federico. 1945. *Gramática histórica de la lengua castellana*. Buenos Aires: El Ateneo.

Haspelmath, Martin. 2011. "The Gradual Coalescence into 'Words' in Grammaticalization." In *The Oxford Handbook of Grammaticalization*, ed. by Bernd Heine and Heiko Narrog, 342–355. Oxford: Oxford University Press.

Heine, Benrd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511613463

Langacker, Ronald. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*, Vol 1. Stanford, CA: Stanford University Press.

Lehmann, Christian. 2002. "New Reflections on Grammaticalization and Lexicalization." In *New Reflections on Grammaticalization*, ed. by Ilse Wischer and Gabriele Diewald, 1–18. Amsterdam: John Benjamins. DOI: 10.1075/tsl.49.03leh

Maldonado, Ricardo. 1999. *A media voz: problemas conecptuales del clítico* se. México: Universidad Nacional Autónoma de México.

Myhill, John. 1995. "Change and Continuity in the Functions of the American English Modals." *Linguistics* 33(2): 157–211. DOI: 10.1515/ling.1995.33.2.157

Poplack, Shana. 2001. "Variability, Frequency and Productivity in the Irrealis Domain of French." In *Frequency and the Emergence of Linguistic Structure*, ed. by Joan Bybee and Paul Hopper, 405–428. Amsterdam: Benjamins. DOI: 10.1075/tsl.45.20pop

Rice, Sally and Kaori Kabata. 2007. "Crosslinguistic Grammaticalization Patterns of the allative." *Linguistic Typology* 11: 451–514. DOI: 10.1515/LINGTY.2007.031

Riiho, Timo. 1979. *Por y para. Estudio sobre los orígenes y la evolución de una oposición prepositiva iberorrománica*. Helsinki: Societas Scientiarum Fennica.

Solalinde, Antonio. 1930. *Alfonso el Sabio: General Estoria, primera parte*. Madrid: Centro de Estudios Históricos.

Torres Cacoullos, Rena. 2006. "Relative Frequency in the Grammaticization of Collocations: Nominal to Concessive *a pesar de*." In *Selected Proceedings of the* 8th *Hispanic Linguistics Symposium*, ed. by Timothy A. Face and Carol E. Klee, 37–49. Somerville: Cascadilla Proceedings Project.

Torres Cacoullos, Rena and Joseph Bauman. 2014. (in press). "Preposiciones III: Por, pora, para." *Sintaxis histórica de la lengua española*, part III: *Adverbios, preposiciones y conjunciones*, Concepción Company Company (director). Mexico: Fondo de Cultura Económica y Universidad Nacional Autónoma de México.

Torres Cacoullos, Rena and Scott A. Schwenter. 2005. "Towards an Operational Notion of Subjectification." In *Proceedings of the 31st Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Prosodic Variation and Change*, ed. by Rebecca T. Cover and Yuni Kim, 347–358. Berkeley, CA: Berkeley Linguistics Society.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. "Empirical Foundations for a Theory of Language Change. In *Directions for Historical Linguistics*, ed. by Winfred P. Lehmann and Yakov Malkiel, 95–195. Austin, TX: University of Texas Press.