

Author accepted manuscript

The future is in the past:

A diachronic analysis of variable future-in-the-past expression in Spanish

Danielle Daidone and Sara Zahler

Indiana University

Note: This article is under copyright and the publisher should be contacted for permission to re-use or reprint the material in any form. Please cite the final published version.

Daidone, D. & Zahler, S. (2016). The future is in the past: A diachronic analysis of variable future-in-the-past expression in Spanish. In A. Cuza, L. Czerwionka, & D. Olson (Eds.), *Inquiries in Hispanic linguistics: From theory to empirical evidence* (pp. 317-334). Philadelphia, PA: John Benjamins. <http://dx.doi.org/10.1075/ihll.12.17dai>

Abstract:

This study examines variation in Spanish future-in-the-past expression across six centuries in order to determine if this variation parallels canonical future variation diachronically. Future-in-the-past tokens in the conditional (e.g., *iría* “would go”) and imperfect *go*-periphrasis (e.g., *iba a ir* “was going to go”) were extracted from two corpora and coded according to factors known to condition the variation between synthetic future (e.g., *irá* “will go”) and periphrastic future (e.g., *va a ir* “is going to go”). Results indicate a rise in the use of the periphrastic form, analogous to the rise in the use of the periphrastic future. Additionally, most factors previously reported to influence canonical future variation were found to play a comparable role in future-in-the-past variation over time.

Keywords: grammaticalization, future-in-the-past, diachronic variation, Spanish

1. Introduction

In Spanish, the three most common forms used to denote subordinate clause events that are subsequent to those in their corresponding main clauses (also known as future-in-the-past) have been shown to be the conditional (1a), the imperfect *go*-periphrasis (1b), and the imperfect (1c) (Radanova-Kusceva & Kitova-Vasileva, 1985; Sarrazin & Azzopardi, 2012).

- (1) a. *Les dió licencia y esperanzas de que presto iría a Judea*
(CORDE, 1580-1630).
'He gave them permission and hope that soon he would go to Judea.'
- b. *Mañana hemos dicho que íbamos a ir a la Sierra*
(CREA, 1980-2004 oral).
'We've said that tomorrow we were going to go to the Sierra.'
- c. *Decía... que el Conde se iba mañana*
(CORDE, 1580-1630).
'He was saying that the Count would go tomorrow.'

Several researchers have claimed that the forms used to express future-in-the-past are similar to the forms that convey canonical future meaning in the Romance languages (e.g., Lozano, 1988; Radanova-Kusceva & Kitova-Vasileva, 1985; Stockwell, Bowen, & Martin, 1965). The connection is also noted in Spanish grammars beginning in the 19th century (e.g., Alarcos Llorach, 1973; Bello, 1869; Di Tullio, 1997). For example, Alarcos Llorach (1973) stated that the conditional shares its "particularities" with the synthetic future form (p. 152).

As such, the conditional form *iría* 'would go' in future-in-the-past reference would be analogous to the synthetic future form *irá* 'will go' in canonical future reference, the imperfect *go*-periphrasis *iba a ir* 'was going to go' would function similarly to its present equivalent *va a ir* 'is going to go,' and the imperfect *iba* 'was going' would be comparable to the futurate use of the present tense *va* 'goes.'

Despite the fact that future-in-the-past forms have been claimed to be analogous to canonical future forms, empirical evidence has yet to verify whether these analogous forms are used the same way across the two contexts. Previous research has shown that the synthetic future, the periphrastic future, and the present can all express the future, and their use is conditioned by various factors. If all of the forms used to express future-in-the-past are analogous to those expressing the canonical future, then we expect them to have the same pattern of variation. However, no one has examined whether this is actually the case. Most previous research solely presents overall rates of forms conveying future-in-the-past or simply states that the future-in-the-past and canonical future forms are analogous. As such, the goal of the current study is to examine the diachronic development of future-in-the-past variation in Spanish through the comparative variationist method. The distributions of variants over time as well as the factors that affect these distributions will be compared with previous diachronic research on the canonical future in order to determine whether the future-in-the-past forms display the same variable patterning as their canonical future counterparts.

2. Previous research on future-in-the-past

The expression of future-in-the-past in Spanish has only been investigated in two studies, which present differing results in terms of variant distribution. The first of these

studies, Radanova-Kusceva and Kitova-Vasileva (1985), examined written texts that spanned the 1940s to 1985. Of the 354 cases of future-in-the-past that they extracted, 67% were the conditional, 23% were the imperfect *go*-periphrasis, 4% were the imperfect indicative, and 6% were comprised of other forms such as the imperfect subjunctive and the periphrastic conditional. As a result, they claimed that the conditional is the main form that has the function of future-in-the-past in Spanish.

On the other hand, Sarrazin and Azzopardi (2012), who looked at the rates of the imperfect *go*-periphrasis and the conditional used for this function in oral data from 1995 to 2004, found that of their 251 cases of future-in-the-past, 76% were the imperfect *go*-periphrasis, while 24% were the conditional, making the *go*-periphrasis the dominant form to express future-in-the-past. The explanation for this discrepancy could be two-fold: first, the corpus that Sarrazin and Azzopardi (2012) employed was more recent than that used by Radanova-Kusceva and Kitova-Vasileva (1985). Therefore, it may be the case that an increase in the *go*-periphrasis occurred between the two time periods, although this scenario is unlikely to account for the entire difference between the two studies, since grammaticalization (as explained in section 3) is a gradual process (Bybee et al., 1994). Another explanation for the discrepancy results from the difference in texts used; in Radanova-Kusceva and Kitova-Vasileva (1985), the authors examined a written corpus of literary texts, while Sarrazin and Azzopardi (2012) searched oral data. Since oral discourse tends to contain more innovations than literary texts, it may be the case that the data Sarrazin and Azzopardi used represent a more advanced step in the grammaticalization process, similar to the difference evidenced in canonical future expression between 20th century written and 20th century oral data in Aaron (2006) and in Poplack and Malvar (2007).

These studies suggest that the distribution of forms used to express future-in-the-past in Spanish has changed over time, with an increase in the frequency of the imperfect *go*-periphrasis. However, this suggestion has yet to be examined as no diachronic studies of the future-in-the-past expression have been undertaken. Furthermore, while both Radanova-Kusceva and Kitova-Vasileva (1985) and Sarrazin and Azzopardi (2012) reported variation in the forms used to express this function, neither examined factors that could condition this variation. Thus, it remains unclear from the previous research on future-in-the-past what function each form has within the future-of-the-past context and whether these roles are similar to their analogous canonical future counterparts.

3. Grammaticalization theory

Grammaticalization theory provides theoretical support for the idea that variation in forms expressing future-in-the-past may be similar to variation of forms in the canonical future context. Grammaticalization is defined as the diachronic process by which lexical items or constructions come to serve grammatical functions, or by which grammatical items develop new grammatical functions (Hopper & Traugott, 1993). One of the tenets of grammaticalization theory is that the original lexico-semantic representation of a form undergoing grammaticalization constrains the grammaticalization path that this form takes; thus forms undergoing grammaticalization that derive from similar lexico-semantic sources tend to follow similar paths (Bybee, Perkins, & Pagliuca, 1994). Since the periphrastic forms used to express future-in-the-past and canonical future both utilize the verb *ir* ‘to go,’ they are expected to develop similarly over time and be favored in the same contexts. Likewise, since the canonical synthetic future form and the future-in-the-past conditional form both developed from *infinitive + conjugated haber* ‘to have’ periphrases (*haber* was conjugated in the present for the canonical future form and in the imperfect for the conditional form)

(Bybee et al., 1994), grammaticalization theory predicts that they will develop similar meanings and functions across time. The imperfect and the present did not grammaticalize from periphrases in Latin but rather descend directly from their corresponding Latin tenses (Pharies, 2007), and as such, it is difficult to predict their behavior based on grammaticalization theory.

4. Previous research on the canonical future

Grammaticalization theory, as well as Spanish grammars and previous research in Romance, suggest that future-in-the-past and canonical future expression are analogous. In order to study whether the factors that constrain canonical future variation also constrain future-in-the-past variation, it is first necessary to determine what factors influence canonical future expression. Thus, in this section we turn to research on the canonical future. Although there exist numerous synchronic studies of the future in different dialects of Spanish (e.g., Blas Arroyo, 2008; Lastra & Butragueño, 2010; Orozco, 2005), for the purposes of the current study we focus on the previous research that employed variationist methodologies and examined the diachronic development of future forms.

Aaron (2006) examined the variation in Spanish between the synthetic future and periphrastic future forms across time, drawing on Peninsular Spanish documents from the mid-13th through early 21st centuries. She found that the *go*-periphrasis was virtually non-existent in the corpus until the 17th-18th century, at which point it constituted just 4% of future forms. This rate increased to 13% in the 19th century data and 25% in the late 20th and early 21st century written data. In the 20th century oral data, the periphrastic form accounted for the majority of instances of future expression, i.e., 59% of tokens. Overall, Aaron's (2006) results reveal that the use of the periphrastic future has gradually increased over time, while the synthetic future has been losing ground.

In terms of the variables found to condition this variation, Aaron (2006) demonstrated that verb class was the most consistent factor determining the use of the periphrastic future in Spanish, as it was significant in all centuries for which a variable rule analysis was conducted (17th, 19th, 20th/21st written, and 20th spoken). In her data, motion verbs and dynamic non-motion verbs consistently favored the periphrastic future, while stative, perception, and psychological verbs disfavored it.

Adverbial modification has also been found to condition the diachronic variation in canonical future expression. In Aaron's (2006) results, the absence of a temporal modifier favored the periphrastic future in all centuries except for the 19th, in which this factor was not significant. Both specific and non-specific temporal adverbials disfavored the periphrastic future, with non-specific adverbials the least favorable context for this variant.

While sentence polarity has been reported to significantly influence variation in future expression, its effect has not been consistent over time. In the 17th century data from Aaron (2006), affirmative sentences slightly favored the periphrastic future while negative polarity highly disfavored it. Polarity was not significant in the 19th century and 20th/21st century written data, but it re-emerged as significant in the 20th century oral data with the opposite direction of effect: in modern data, negative sentences favored the periphrastic form.

Furthermore, sentence modality has been reported to have a significant effect on canonical future variation. In Aaron (2006) sentence type was not significant in earlier centuries (17th and 19th) but was significant in both 20th century corpora; interrogative sentences favored the periphrastic future while declaratives slightly disfavored it.

Whether the future form occurred in a main or subordinate clause was found to be important for variation in future forms as well. Clause type was only significant in the 19th century data, in which subordinate clauses favored the periphrastic form (Aaron, 2006).

Lastly, although the effect of temporal distance on diachronic canonical future variation has not been investigated in Spanish, it was found to condition such variation in Portuguese and French. Poplack and Malvar (2007) found evidence that proximal events favored the periphrastic future and disfavored the synthetic future in their 19th century Portuguese data, although this effect disappeared by the 20th century. Similarly, Poplack and Dion (2009) reported that reference to distal events favored the synthetic future in French in the 19th century, but there was no effect of temporal distance in their 20th century results. This tendency was also true for the results of synchronic studies on canonical future variation in modern Spanish, which have generally found that proximate events favor the periphrastic future while distant or indefinite temporal distance favor the synthetic future (e.g., Blas Arroyo, 2008; Lastra & Butragueño, 2010; Orozco, 2005).

Overall, this research shows that in Spanish the use of the synthetic form has decreased over time, whereas the presence of the periphrastic form has increased. These studies also evidenced a multitude of factors that condition the variation in canonical future expression over time. While clause type is not relevant to the current study because future-in-the-past readings are primarily limited to subordinate clauses, all of the other variables significant in these previous studies – verb class, adverbial modification, polarity, sentence modality, and temporal distance – are candidates for the factors that may play a role in the variation of future-in-the-past expression and can serve as a point of comparison with the canonical future.

5. The current study

The purpose of the current study is to characterize the distribution and variation in future-in-the-past forms over time and compare this to previous diachronic research on the canonical future in Spanish. Originally we intended to examine the variation between the imperfect, conditional, and imperfect *go*-periphrasis, as these are the three most frequent forms reported to express future-in-the-past and are analogous to the futurate present, synthetic future, and periphrastic future, respectively, in canonical future reference. However, while extracting imperfect tokens with a future-in-the-past interpretation we found that they were extremely infrequent; only one out of the first 198 tokens extracted of imperfect in subordinate clauses actually expressed future-in-the-past. Given this low rate of relevant imperfect forms in data from the earliest time period we examined (1580-1630), and the fact that Sarrazin and Azzopardi (2012) reported that the imperfect only accounted for 4% of future-in-the-past tokens in the late 20th and early 21st century, we decided to restrict the present investigation to conditional and periphrastic forms. Additionally, this allows for a more direct comparison with Aaron (2006), who examined future variation in terms of the synthetic and periphrastic future and did not include the futurate present. Thus, the revised research questions that guide this study are the following:

1a) What is the distribution of the conditional and periphrastic forms when expressing future-in-the-past in Spanish across four data sets representing four time periods (1580-1630, 1780-1830, 1980-2004 written, and 1980-2004 oral)?

1b) How does the change in the distribution of forms expressing future-in-the-past compare to that of the canonical future?

2a) What linguistic factors affect the variation between the conditional and periphrastic forms across the four time periods?

2b) How does the influence of these linguistic factors over time compare to that of the canonical future?

In terms of predictions, we expect to find an increase in the periphrastic form which mirrors that found for the canonical future context in previous research. Based on grammaticalization theory, we also hypothesize that the linguistic factors that constrain canonical future variation will have a comparable effect on forms expressing future-in-the-past.

6. Method

6.1 Corpora

The data for this study come from two corpora from the *Real Academia Española* (2008): the *Corpus Diacrónico del Español* (CORDE), which is a diachronic database consisting of texts from as early as the year 800 until 1975, and the *Corpus de Referencia del Español Actual* (CREA), which contains data from 1975 until 2004. We limited our search within these corpora to data solely from Peninsular Spanish for two reasons: first, because data from other dialect regions would not be well represented until more recent time periods, and secondly, for better comparability with Aaron's (2006) study which examined diachronic canonical future variation in Peninsular Spanish. Our searches in the corpora were restricted to the genres of theater and narrative for CORDE and theater and oral for CREA; this was done in order to have a sample as close to spoken data and as consistent across the centuries as possible. Three specific periods of time were examined: 1580-1630, 1780-1830, and 1980-2004, between each of which there is a span of 150 years. These time frames were chosen in increments similar to Aaron (2006), who used diachronic increments of approximately 150-200 years for the following reason:

This time period was considered sufficient to observe change in progress in written language; a lapse of only 50 years would most likely be too short to note any significant systematic shift in patterns, while a lapse of 300 years would most likely leave intermediate steps in change out of the picture. (p. 34)

As such, we chose 150 year installments for comparability with Aaron (2006), and in order to capture possible changes in a way that allows us to observe discernable differences between periods without missing important changes. Lastly, we divided the 1980-2004 data set into written (theater) and oral data for comparability to Aaron (2006) and Poplack and Malvar (2007), and because written discourse is often a poor approximation of speech since it does not reflect the most current linguistic developments that first occur in spoken language (Van Herk & Poplack, 2003).

6.2 Variables

6.2.1 Independent variables

Based on the factors that were found to significantly condition variation in previous diachronic research on canonical future expression, the following independent linguistic variables were considered in this study: verb class, temporal modification, polarity, sentence

modality, and temporal distance. Information on the operationalization of these factors is summarized in Table 1 and explained in the subsequent paragraphs.

Table 1

Independent variables

Variables	Categories
1. Verb class	Motion: <i>ir</i> <i>llegar</i> <i>correr</i> <i>caer</i> Dynamic non-motion: <i>hacer</i> <i>dar</i> <i>sacar</i> <i>recibir</i> Psychological : <i>querer</i> <i>saber</i> <i>entender</i> <i>recordar</i> Stative: <i>ser</i> <i>estar</i> <i>resultar</i> <i>corresponder</i>
2. Temporal modification	None Specific Non-specific
3. Polarity	Affirmative Negative
4. Sentence modality	Declarative Interrogative
5. Temporal distance	More than a year Within a year (proximate) Indeterminate Unknown

Verb class. Future-in-the-past forms of sixteen specific verbs were extracted from the corpora (Table 1). These sixteen verbs were equally distributed across four verb classes: motion, dynamic non-motion, psychological, and stative. In order to control for any possible frequency effects, two high frequency and two moderate frequency verbs were chosen within each of the four classes. High frequency was specified as those verbs within the top 100 most frequent words, while moderate frequency was specified as those verbs within the top 200-600 most frequent words according to a frequency dictionary of Spanish (Davies, 2006).

Temporal modification. If temporal modification was present in the clause containing the future-in-the-past verb, it was coded as specific or non-specific; otherwise, temporal modification was coded as none. Specific temporal modification indicated that the future-in-the-past event was to occur either at a specific moment in time, e.g., *Me daba el corazón que hoy iba a recibir un regalo muy, muy, pero que muy lindo* ‘My heart told me that today I was going to receive a very, very nice gift’ (CREA, 1980-2004 oral), or at the moment of the completion of another specific event, such as *Siempre esperé confiada en que llegaría el momento de tu tranquilidad cuando editases tu trabajo* ‘I always confidently expected that the moment of your tranquility would arrive when you published your work’ (CREA, 1980-2004 written). Temporal modification was considered non-specific when the exact moment of the future-in-the-past action was not indicated, such as in the example *Pero también*

estábamos seguros de que tarde o temprano lo entenderías ‘But we were also sure that sooner or later you would understand it’ (CREA, 1980-2004 written) or when the temporal modification indicated that the future-in-the-past action was not going to occur, as in *Hermano creía que nunca iba a llegar el momento* ‘Brother believed that the moment was never going to arrive’ (CREA, 1980-2004 written).

Polarity. Clauses containing future-in-the-past forms were coded as either affirmative, e.g., *consolóme con decirme que su vuelta sería con brevedad* ‘He consoled me by telling me that his return would be brief’ (CORDE, 1580-1630), or negative, e.g., *y les protestó que jamás haría cosa alguna que pudiese desagradarlos* ‘and he promised them that never would he do something that could upset them’ (CORDE, 1780-1830).

Sentence modality. Sentences in which future-in-the-past forms appeared were coded as either declarative, e.g., *Y él me juraba cada año que al siguiente, que al siguiente nos iríamos...* ‘And he swore to me each year that during the next one, during the next one we would go...’ (CREA, 1980-2004 written), or interrogative, e.g., *¿...A qué hora dijo que llegaría?* ‘At what time did he say he would arrive?’ (CREA, 1980-2004 written).

Temporal distance. This variable coded the temporal distance between the future-in-the-past action and the speech event. This factor was determined by the context surrounding the token, both within the sentence itself as well as in surrounding sentences. Temporal distance was determined based on whether an event was referred to as occurring within a year or after a year from the event in the main clause. These time frames were normally ascertained by the presence of temporal modifiers such as *pronto* ‘soon,’ *antes de veinticuatro horas* ‘within twenty-four hours,’ *en menos de quince días* ‘in less than fifteen days,’ or *en seis meses* ‘in six months.’ A token was coded as ‘indeterminate’ in a case such as *Pero también estábamos seguros de que tarde o temprano lo entenderías* ‘But we were also sure that sooner or later you would understand it’ (CREA, 1980-2004 written), in which a vague or changeable time frame was given. ‘Unknown’ includes future-in-the-past actions for which it is impossible to determine the distance to the speech event based on the context, such as with the sentence *Sabía que iban a llegar las naves* ‘He knew that the ships were going to arrive’ (CREA, 1980-2004 written).

6.2.2 Dependent variable

In CORDE and CREA for the time periods and genres mentioned above, we conducted a search for tokens of the conditional (e.g., *daría* ‘I would give’) and imperfect *go*-periphrasis (e.g., *iba a dar* ‘I was going to give’) for the 16 verbs specified in Table 1 in subordinate clauses. All persons and number combinations were searched, including well-known alternate spellings such as *iva*, *yba* and *yva* for *iba* ‘he was going’ in the 1580-1630 time period. All tokens that referred to a future-in-the-past event were then extracted and coded according to the linguistic variables discussed in the preceding section. We excluded conditional or epistemic uses of the conditional form, e.g., *Hace cinco años que empezó, haría en junio cinco años que empezó* ‘Five years ago it started, it must have been five years ago in June that it started’ (CREA, 1980-2004 oral) as well as purely allative uses of the *go*-periphrasis, e.g., *Pues si yo no le trujera, se iba a una ermita a hacer penitencia, y si no la hallaba hecha, que él la haría* ‘Well if I didn’t bring him, he would go to a chapel to do penance, and if he didn’t find it done, he would do it’ (CORDE, 1580-1630), since neither of these uses encodes a future-in-the-past reading.

6.3 Analysis

First, the distribution of the imperfect *go*-periphrasis and the conditional form was determined within each time period, and a Pearson chi-square analysis was performed in SPSS 22 in order to determine if the difference in the distribution of the variants was significant across time periods. Subsequently, Pearson chi-square analyses were performed on the distribution of the periphrastic and conditional forms for each factor within each time period. These chi-square analyses indicate whether the distribution of the future-in-the-past forms is significantly different across categories (e.g., affirmative versus negative) within a particular factor (e.g., polarity).

7. Results

Overall, the rate of the imperfect *go*-periphrasis increased across time, $\chi^2(3, N = 934) = 307.12, p < .001$, as presented in Table 2. This increase is demonstrated in the token count and % *go*-periphrasis columns which demonstrate the proportion of *go*-periphrasis to the total token count from each time period. Given that the token counts constitute the total number of periphrastic tokens and conditional tokens used to express future-in-the-past in each time period, the rise in the use of the periphrastic form also indicates an equivalent decrease in the use of the conditional within the future-in-the-past context diachronically.

Rate of imperfect go-periphrasis across time periods

Time period	Token count	% <i>go</i> -periphrasis
1580-1630	35/473	7.4%
1780-1830	13/113	11.5%
1980-2004 written	48/179	26.8%
1980-2004 oral	123/169	72.8%

A Pearson chi-square test indicated that the proportion of imperfect *go*-periphrasis to the conditional is not significantly different between 1580-1630 and 1780-1830 ($p > .05$). However, the distribution of the periphrastic and conditional variants was significantly different in the 1980-2004 written data compared to the 1580-1630, 1780-1830, and 1980-2004 oral data ($p < .05$). Similarly, the distribution of the variants in the 1980-2004 oral data was significantly different from all other data sets ($p < .05$). These results suggest that the change in rate of the periphrastic variant, while not significant between 1580-1630 and 1780-1830, increased significantly between 1780-1830 and the 1980-2004 data, with a significantly greater use of the periphrastic in oral data compared to written data in the 1980-2004 period.

7.1 Temporal modification

In Table 3, the distribution of the imperfect *go*-periphrasis according to temporal modification is given for each time period. A Pearson chi-square test indicates that within time periods the distribution of the two forms by adverbial modification was statistically significant for 1580-1630 only, $\chi^2(2, N = 473) = 7.81, p = .020$. In this time period, the *go*-periphrasis occurred significantly more often when there was no temporal modification since it did not occur with any temporal modifiers. On the other hand, distribution of the variants was not significant according to temporal modification in any of the other time periods (1780-1830: $\chi^2(2, N = 113) = 3.01, p = .222$; 1980-2004 written: $\chi^2(2, N = 179) = 2.51, p =$

.286; 1980-2004 oral: $\chi^2(2, N = 169) = 3.12, p = .210$). Nevertheless, the distribution of variants in these three time periods is noteworthy. Other than one token of a specific adverbial in 1780-1830, the *go*-periphrasis did not occur with temporal modification until the 1980-2004 oral and written data, where it occurred with all three types: no temporal modification, specific modifiers, and to a lesser extent non-specific modifiers.

Table 3

Distribution of the periphrastic form by temporal modification

Temporal modification	1580-1630		1780-1830		1980-2004 written		1980-2004 oral	
	N	%	N	%	N	%	N	%
None	35/392	8.9	12/83	14.5	39/137	28.5	105/139	75.5
Specific	0/33	0	1/15	6.7	4/11	36.4	7/11	63.6
Non-specific	0/48	0	0/15	0	5/31	16.1	11/19	57.9

Collectively, these results suggest that the periphrastic variant was more strongly constrained by temporal modification earlier on, but was less so in later time periods.

7.2 Temporal distance

The distribution of the imperfect *go*-periphrasis according to temporal distance is displayed in Table 4. This distribution was statistically significant for the 1580-1630 time period ($\chi^2(2, N = 473) = 8.04, p = .018$) and the 1980-2004 oral data ($\chi^2(2, N = 169) = 14.50, p = .001$). In the first time period, the periphrastic variant only occurred with unknown temporal reference. On the other hand, in the 1980-2004 oral data the periphrastic variant occurred with all types of temporal reference and more so with proximate and unknown temporal reference than with indeterminate or more than a year reference. Although not significant, the distribution of variants in the 1780-1830 time period was very similar to that of the 1580-1630 time period; other than one token of the periphrastic variant with indeterminate reference, the periphrastic variant only occurred with unknown temporal reference. Similarly, although the distribution by temporal distance was not significant for the 1980-2004 written data set, the use of variants with different types of temporal distance was comparable to the 1980-2004 oral data set; the periphrastic variant occurred most often with proximate reference, then unknown, and least with indeterminate or distant temporal reference.

Table 4

Distribution of the periphrastic form by temporal distance

Temporal distance	1580-1630		1780-1830		1980-2004 written		1980-2004 oral	
	N	%	N	%	N	%	N	%
Unknown	35/390	9.0	12/87	13.8	37/135	27.4	101/134	75.4
Proximate (within a year)	0/34	0	0/12	0	4/10	40.0	14/15	93.3
Indeterminate/ Year +	0/49	0	1/14	7.1	7/34	20.6	8/20	40.0

Across time periods, the results for temporal distance are similar to those for adverbial specification in that the imperfect *go*-periphrasis was restricted mainly to unknown temporal reference in the first two time periods, but by the 1980-2004 written and oral data, the

periphrastic form was used with all types of temporal distance, appearing most often with proximate reference.

7.3 Polarity.

The distribution of the variants in affirmative and negative contexts was statistically significant for the 1980-2004 written data, $\chi^2(1, N = 179) = 8.612, p = .003$. In this time period, the periphrastic variant occurred significantly more in negative contexts than in affirmative ones. This distribution represents a difference from the 1580-1630 data ($\chi^2(1, N = 473) = 2.74, p = .098$) and the 1780-1830 data ($\chi^2(1, N = 113) = .97, p = .325$), in which, although not statistically significant, the periphrastic variant did not occur in any negative contexts. This distribution also differs from the 1980-2004 oral data, in which the periphrastic variant occurred at fairly similar rates in both affirmative and negative contexts and thus polarity did not have a significant effect ($\chi^2(1, N = 169) = .88, p = .349$).

Table 5
Distribution of the periphrastic form by polarity

Polarity	1580-1630		1780-1830		1980-2004 written		1980-2004 oral	
	N	%	N	%	N	%	N	%
Affirmative	35/441	7.9	13/106	12.3	40/166	24.1	117/159	73.6
Negative	0/32	0	0/7	0	8/13	61.5	6/10	60.0

Considering the change in distribution across time periods, it appears that the imperfect *go*-periphrasis was highly constrained in the 1580-1630 and 1780-1830 time periods, occurring only in affirmative contexts. However, by the late 20th/21st century, in both written and oral data, the periphrastic variant occurred in both affirmative and negative contexts, though it was favored in negative contexts in the 1980-2004 written data.

7.4 Sentence modality

Other than in the 1780-1830 data, the periphrastic form occurred in at least one interrogative sentence per time period. However, since there were never more than four tokens of interrogative sentences in any given time period, it was not possible to ascertain any pattern of variation in terms of sentence modality.

7.5 Verb class

According to a Pearson chi-square test, the proportion of the periphrastic variant according to verb class was statistically significant for 1580-1630 ($\chi^2(1, N = 473) = 5.41, p = .019$), 1980-2004 written ($\chi^2(1, N = 179) = 13.20, p < .001$), and 1980-2004 oral data ($\chi^2(1, N = 169) = 12.92, p < .001$), but not for 1780-1830 ($\chi^2(1, N = 113) = 2.07, p = .150$). This lack of significance in the 1780-1830 time period may in part be due to the low token count in this data set.

Table 7

Distribution of the periphrastic form by verb class

Verb class	1580-1630		1780-1830		1980-2004 written		1980-2004 oral	
	N	%	N	%	N	%	N	%
Dynamic non-motion	26/262	9.9	9/57	15.8	25/56	44.6	49/54	90.7
Psychological, Stative, Motion	9/211	4.3	4/56	7.1	23/123	18.7	74/115	64.3

In each time period where the distribution was significant, the periphrastic variant occurred more frequently with dynamic non-motion verbs than with psychological, stative, and motion verbs. These results suggest that the effect of verb class has remained consistent across time.

8. Discussion

In terms of the distribution of future-in-the-past forms in our Peninsular Spanish data, the rate of the imperfect *go*-periphrasis significantly increased across time periods at the expense of the conditional. While the imperfect *go*-periphrasis rarely appeared in the 1580-1630 and 1780-1830 data, its rate of occurrence increased in the 20th/21st century written data and in fact exceeded the use of the conditional in the 20th/21st century oral data. Comparing this distribution to that of forms expressing the canonical future in Peninsular Spanish, the results are quite similar. Aaron (2006) found 25% *go*-periphrasis in her 20th/21st century written data and 59% in her 20th century oral data. This rise in the 20th century of the periphrastic future was also evidenced in Brazilian Portuguese (Poplack & Malvar, 2007) and Canadian French (Poplack & Dion, 2009).

In our examination of the linguistic factors that affected the variation between conditional and periphrastic forms, we found that the effect of sentence modality and subordinate verb class generally remained consistent across time. Sentence modality was not significant in any of the time periods, which is most likely due to the lack of interrogative contexts within each time period. This result differs from research on canonical future variation in Spanish, as Aaron (2006) found that the periphrastic future form was favored in interrogatives in her 20th/21st century written and 20th century oral data. However, it is difficult to make a comparison due to the lack of interrogatives tokens in our data. On the other hand, in our study, verb class was significant in all time periods except for 1780-1830. In each data set, the periphrastic variant occurred more frequently with dynamic non-motion verbs than with stative, motion, or psychological verbs. This consistent effect of verb class in our data largely mirrors that of Aaron (2006), who found that the periphrastic form was favored by dynamic non-motion verbs across time and disfavored by stative, psychological, and perception verbs. The only difference between our results and Aaron's (2006) is that in her data, across all time periods, motion verbs also favored the periphrastic variant, while in our data, the *go*-periphrasis occurred less frequently with motion verbs. However, Aaron (2006) excluded the verb *ir* 'to go' from her category of motion verbs, and we did not. This difference may likely be the root of our differing results, since Poplack and Malvar (2007) also included *ir* 'to go' in their category of motion verbs, and found as well that motion verbs disfavored the periphrastic future variant in their Portuguese data.

The effect of temporal modification, temporal distance, and polarity changed across time periods. For these three factors, the results suggest that the periphrastic variant was

highly constrained in the first two time periods, 1580-1630 and 1780-1830, since it almost categorically did not occur with any temporal modification, with known temporal distance, or in negative contexts. However, by the 1980-2004 time period for both written and oral data, the imperfect *go*-periphrasis occurred with all categories of each factor, indicating that its use was less constrained. Nevertheless, its distribution amongst the categories of each of these factors differed, pointing to the continued influence of temporal modification, temporal distance, and polarity on future-in-the-past variation.

For temporal modification, the periphrastic variant spread to specific and non-specific modifiers in the 20th/21st century written and oral data, occurring more with specific than non-specific modifiers. These results mirror the distribution found in Aaron's (2006) data. In her study, in the 17th century, the periphrastic future was highly disfavored by any temporal modification, but it spread to all types, both specific and non-specific, by the 20th century, although it remained disfavored with specific modifiers and even more so with non-specific ones.

In terms of temporal distance, the periphrastic variant in our data spread to all types of temporal distances in the 20th/21st century written and oral data, but it occurred more frequently in proximate contexts and less frequently with distal or indeterminate temporal reference. Aaron (2006) did not examine temporal distance for Spanish; however, research on Portuguese (Poplack & Malvar, 2007) and on Canadian French (Poplack & Dion, 2009) found more periphrastic future with proximate temporal reference and less with distal temporal reference in the 19th century. By the 20th century, this effect had disappeared in Portuguese, and had reversed in French but with a small effect size. Nevertheless, synchronic contemporary research on future variation in Spanish indicates that the periphrastic variant is typically used more in proximate contexts, while the synthetic variant generally is used more in distal contexts (e.g., Blas Arroyo, 2008; Lastra & Butragueño, 2010; Orozco, 2005). Thus, while comparisons for our two earliest data sets are not possible due to the lack of diachronic research on the effect of this factor on future variation, our results appear to corroborate those found for contemporary future variation.

Lastly, the distribution of variants according to polarity is similar to that of Aaron (2006) for our first three time periods. In the 1580-1630 and 1780-1830 data sets, the *go*-periphrasis did not occur with any negative tokens. In the 20th/21st century written data, the periphrastic form occurred more frequently in negative contexts, and this distribution was significant. In Aaron's (2006) data, the periphrastic form was at first disfavored in negative contexts, but by the 20th century oral data, it was favored in negative contexts. However, our results differed from hers in that the distribution of variants according to polarity in our 20th/21st century oral data is not significant, and there no longer seems to be an effect of polarity. This may be due to differences between the variation of future-in-the-past forms and that of canonical future forms, or it may be due to the small token count of negative contexts in our 20th/21st century oral data.

Overall, it is not surprising that the factors that constrain the variation in forms expressing future-in-the-past and those expressing canonical future are similar. Although these forms occur in different contexts, they have the same lexico-semantic origins as their counterparts. Specifically, for the *go*-periphrases, the forms originally expressed movement towards a goal, which developed meanings of proximate intentionality and later futurity. As this form moved from an ongoing-spatial meaning to a future-temporal meaning, the time frame within which these future actions occurred expanded outwards (Bybee et al., 1994). This is likely why the *go*-periphrases first occurred with no temporal modification, then specific and proximate adverbials, expressing certain proximate future events, and lastly non-specific and distant adverbials. Additionally, since the original meaning of the variant was that of motion towards a goal, the verbs that it originally co-occurred with were dynamic non-

motion verbs since the speaker was originally physically going to another place in order to perform an action. As such, it is not surprising that this form is disfavored by stative or psychological verbs. On the other hand, since the conditional and synthetic future are the older forms in their contexts and grammaticalized earlier than the periphrastic future (Bybee et al., 1994), they have largely lost the previous constraints on their use that likely existed when they first formed. As such, in the 1580-1630 data, the synthetic form for the future and future-in-the-past context could occur in all contexts and were not highly restricted by the variables examined. However, as the periphrastic variant became the dominant form in the 20th/21st century oral data, the synthetic form became more restricted to the contexts in which the periphrastic variant was the last to enter, i.e., contexts expressing non-specific, distant, or uncertain temporal reference.

9. Conclusion

There are several similarities between our results on diachronic variation in forms expressing future-in-the-past in Peninsular Spanish and the variation reported for canonical future forms diachronically. First, with the exception of motion verbs, verb class seems to have a similar effect on variation in both the future-in-the-past and canonical future contexts, and this effect remains consistent across time. Similarly, the distribution of variants according to temporal modification developed similarly in both contexts. The effect of polarity was also largely similar until the 20th/21st century oral data. Additionally, our 20th/21st century written and oral data showed similar effects of temporal distance as those evidenced in synchronic contemporary studies of future variation. While sentence modality behaved differently in that it was not significant in our study, this lack of significance is most likely due to the low number of interrogative contexts.

Therefore, it seems that, in many respects, the same factors play a role in future-in-the-past variation across time as in canonical future variation. This similarity is also seen in the analogous increase of the periphrastic form within each context, which is greatest between the 20th/21st century written and oral data sets. This resemblance between the variation in each context is perhaps not entirely surprising if one considers the observation that the development of grammaticalizing forms is often constricted by their lexico-semantic origins. Thus, although forms expressing the future and those expressing the future-in-the-past occur in different contexts, the analogous forms within each context appear to be following a similar route of grammaticalization partially determined by their lexico-semantic origins. This mutual evolution between the forms in these two contexts, over 400 years, demonstrates the interconnectedness of the entire tense, mood, and aspect system in Spanish; changes in one variable context can manifest in other variable contexts. As such, the system is a dynamic, linked structure that should be examined from a holistic perspective.

References

- Aaron, J. E. (2006). *Variation and change in Spanish future temporal expression: Rates, constraints, and grammaticization* (Unpublished doctoral dissertation). University of New Mexico, Albuquerque.
- Alarcos Llorach, E. (1973). *Gramática de la lengua española*. Madrid, Spain: Espasa Calpe.
- Bello, A. (1869). *Gramática de la lengua castellana destinada al uso de los americanos*. Bogotá, Colombia: Reimpresión Echeverría.
- Blas Arroyo, J. L. (2008). The variable expression of future tense in Peninsular Spanish: The present (and future) of inflectional forms in the Spanish spoken in a bilingual region. *Language Variation and Change*, 20, 85-126. doi:10.1017/S095439450800001X
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago, IL: University of Chicago Press.
- Davies, M. (2006). *A frequency dictionary of Spanish: Core vocabulary for learners*. New York, NY: Routledge.
- Di Tullio, A. (1997). *Manual de gramática del español*. Buenos Aires, Argentina: Edicial.
- Hopper, P. & Traugott, E. (1993). *Grammaticalization*. Cambridge, England: Cambridge University Press.
- Lastra, Y., & Butragueño, P. (2010). Futuro perifrástico y futuro morfológico en el corpus sociolingüístico de la ciudad de México. *Oralia*, 13, 145-172.
- Lozano, A. G. (1988). The semantics of the Spanish conditional in discourse. *Hispania*, 71, 675-680. doi:10.2307/342931
- Orozco, R. (2005). Distribution of future time forms in Northern Colombian Spanish. In *Selected Proceedings of the 7th Hispanic Linguistics Symposium* (pp. 56-65). Somerville, MA: Cascadilla Proceedings Project.
- Pharies, D. A. (2007). *A brief history of the Spanish language*. Chicago, IL: University of Chicago Press. doi:10.7208/chicago/9780226666846.001.0001
- Poplack, S., & Dion, N. (2009). Prescription vs. praxis: The evolution of future temporal reference in French. *Language*, 85, 557-587. doi:10.1353/lan.0.0149
- Poplack, S., & Malvar, E. (2007). Elucidating the transition period in linguistic change: The expression of the future in Brazilian Portuguese. *Probus*, 19, 121-169. doi:10.1515/PROBUS.2007.005
- Radanova-Kusceva, N., & Kitova-Vasileva, M. (1985). Forme esprimimenti un'azione futura rispetto a un momento di riferimento al passato nell'italiano e nello spagnolo contemporaneo [Forms expressing future action with a reference point in the past in contemporary Italian and Spanish]. *Rassegna Italiana di Linguistica Applicata*, 17(1), 1-31.
- Real Academia Española. (2008). *Corpus diacrónico del español* (CORDE) [online corpus]. Retrieved from <<http://www.rae.es>> [April 16, 2014].
- Real Academia Española. (2008). *Corpus de referencia del español actual* (CREA) [online corpus]. Retrieved from <<http://www.rae.es>> [April 16, 2014].
- Sarrazin, S., & Azzopardi, S. (2012). L'alternance du conditionnel et de la périphrase *itive* à l'imparfait dans des corpus oraux espagnols et français [The alternation of conditional and imperfect *itive* periphrasis in Spanish and French spoken corpora]. *Studia Universitatis Babeş-Bolyai Philologia*, 57(3), 57-68.
- Stockwell, R. P. J., Bowen, D., & Martin, J. W. (1965). *The grammatical structure of English and Spanish*. Chicago, IL: University of Chicago Press.
- Van Herk, G., & Poplack, S. (2003). Rewriting the past: Bare verbs in the Ottawa repository of early African American correspondence. *Journal of Pidgin and Creole Languages*, 18, 231-266. doi:10.1075/jpcl.18.2.03her