ARTICLE IN PRESS

Journal of Memory and Language xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Agreement attraction in Spanish comprehension

Sol Lago [a,*], Diego E. Shalom [b], Mariano Sigman [b], Ellen F. Lau [a], Colin Phillips [a]

[a] Department of Linguistics, University of Maryland, United States
[b] Laboratorio de Neurociencia Integrativa, University of Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 18 February 2014
revision received 8 February 2015
Available online xxxx

Keywords:
Agreement attraction
Number morphology
Spanish
Sentence comprehension

ABSTRACT

Previous studies have found that English speakers experience attraction effects when comprehending subject–verb agreement, showing eased processing of ungrammatical sentences that contain a syntactically unlicensed but number-matching noun. In four self-paced reading experiments we examine whether attraction effects also occur in Spanish, a language where agreement morphology is richer and functionally more significant. We find that despite having a richer morphology, Spanish speakers show reliable attraction effects in comprehension, and that these effects are strikingly similar to those previously found in English in their magnitude and distributional profile. Further, we use distributional analyses to argue that cue-based memory retrieval is used as an error-driven mechanism in comprehension. We suggest that cross-linguistic similarities in agreement attraction result from speakers deploying repair or error-driven mechanisms uniformly across languages.

Introduction

Languages differ in the degree to which their morphological systems convey formal and conceptual distinctions. One important unanswered question is: how does this variation affect core language processing mechanisms? For example, speakers may rely on morphological cues to different extents depending on the availability of these cues in their native language (MacWhinney, 1987; MacWhinney, Bates, & Kliegl, 1984). This could affect the processing of grammatical relationships such as agreement: agreement errors in comprehension might be more common in languages with an impoverished morphology, and specific challenges might arise for learners of a morphologically richer second language (Jia, Aaronson, & Wu, 2002; Jiang, 2004, 2007; McDonald, 2000).

The current study investigates whether morphological variation affects a mechanism crucial to language comprehension: the ability to retrieve previous information from memory (Caplan & Waters, 2013; Gordon, Hendrick, & Johnson, 2001; Lewis & Vasishth, 2005; McElree, Foraker, & Dyer, 2003). The comprehension of subject–verb agreement is likely to involve retrieval: since subjects and verbs can be separated by an unlimited number of words, when a verb is encountered speakers may need to retrieve the number features of the subject noun from memory in order to license agreement. We examine whether cross-linguistic variation affects how speakers process agreement and we ask: is retrieval implemented uniformly across languages, or does it vary depending on the properties of each language?

This question is addressed by comparing the computation of subject–verb agreement in Spanish and English. In English, number morphology is limited, so word order and syntactic information are the most reliable cues to resolve subject–verb dependencies (MacWhinney et al., 1984; Severens, Jansma, & Hartsuiker, 2008). In contrast, agreement morphology in Spanish is both more available

* Corresponding author at: University of Potsdam, Potsdam Research Institute for Multilingualism, Campus Golm, Haus 2, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany.
E-mail address: marlago@uni-potsdam.de (S. Lago).

http://dx.doi.org/10.1016/j.jml.2015.02.002
0749-596X/

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

and functionally more significant, as illustrated below (agreement morphology bolded):

1a. And yet the animal**s** never gave up hope. And **they** never lost, even for an instant, **their** sense of honor and privilege in being member**s** of Animal Farm. (George Orwell, *Animal Farm*)

1b. Y aún así, **los** animal**es** nunca renunci**aron** a **la** esperanza. Y (**ellos**) nunca perdi**eron**, ni siquiera por **un** instante, **sus** sentido**s** de honor y privilegio por ser miembr**os** de **la** Granja de Animal**es**.

These passages highlight several differences between languages. First, number agreement is more available in Spanish. This is the case in nominal phrases, where nouns and all their modifiers carry agreement information (e.g. "su<u>s</u> sentido<u>s</u>", 'their sense'; "lo<u>s</u> animale<u>s</u>", 'the animals'), while plural number in English is mainly marked on the head noun. Further, in Spanish verbs mark agreement for all syntactic persons, and singular and plural verb forms often differ sharply (e.g. "renunció vs. "renunci<u>aron</u>"; 'gave up.3sg' vs. 'gave up.3pl'). Second, agreement morphology is functionally more important in Spanish. One reason for this is that Spanish has a freer word order, so sentence initial position is not as reliable a cue to subjecthood as in English. Relatedly, Spanish is a null-subject language, and subjects can be omitted in sentences (e.g. "<they> lost"). These two properties make verb morphology the main cue for subject identification, which has given rise to the claim that morphological cues are more reliable than positional information in Spanish (Kail, 1989; MacWhinney, 2001).

In the current study, we compare the processing of agreement in Spanish and English, with the goal of examining whether retrieval is implemented differently in languages that contrast in the richness and availability of morphological cues. Cross-linguistic variation is expected under frameworks like the Competition Model (e.g. MacWhinney, 1987; MacWhinney & Bates, 1989), where more available cues are predicted to be acquired first and to most strongly affect adult processing. If Spanish comprehenders rely more on morphological cues than English comprehenders, they might be less susceptible to agreement errors. In the rest of the Introduction, we summarize previous findings of agreement attraction in English and describe a retrieval mechanism that has been proposed underlie these errors. We then discuss how this mechanism might be used differently in Spanish, and present an overview of four experiments that were carried out to examine the relationship between morphological richness and agreement computations.

Attraction in comprehension

In comprehension, agreement attraction facilitates the processing of ungrammatical sentences (Dillon, Mishler, Sloggett, & Phillips, 2013; Nicol, Forster, & Veres, 1997; Pearlmutter, Garnsey, & Bock, 1999; Tanner, Nicol, & Brehm, 2014; Wagers, Lau, & Phillips, 2009). For example, in the ungrammatical sentence "The key to the cabinet(s) <u>are</u> on the table", comprehenders typically read the words following the plural verb more quickly when there is a noun, called an "attractor", that matches the verb in number ("cabinets"). This facilitated processing has been attributed to cue-based memory retrieval (Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; Wagers et al., 2009). When a verb is encountered, speakers use the syntactic, semantic and morphological cues of the verb to retrieve an appropriate subject from memory. Memory chunks corresponding to preceding words and phrases in the sentence are queried in parallel, and the chunk with the most features matching the cues of the verb is the most likely to be retrieved. In the sentence above, this sometimes results in the incorrect retrieval of "cabinets", which allows comprehenders to license the verb in number and yields facilitated processing.

A key piece of evidence for the retrieval account in comprehension is the finding that number-matching attractors affect processing in ungrammatical, but not in grammatical sentences. This finding was first described by Wagers, Lau, and Phillips (2009, henceforth WL&P), who used relative clause constructions (RCs) where plural attractors did not intervene linearly between the critical subject–verb pair (see also Clifton, Frazier, & Deevy, 1999; Staub, 2009, 2010). WL&P manipulated sentence grammaticality and attractor number using a self-paced reading paradigm, and found that plural attractors (e.g. "musicians") resulted in facilitated processing after the verb in ungrammatical sentences (2c vs. 2d) but made no difference in grammatical sentences (2a vs. 2b):

2a. The musician [who the reviewer <u>praises</u> so highly] will probably win a Grammy.

2b. The **musicians** [who the reviewer <u>praises</u> so highly] will probably win a Grammy.

2c. *The musician [who the reviewer <u>praise</u> so highly] will probably win a Grammy.

2d. *The **musicians** [who the reviewer <u>praise</u> so highly will probably win a Grammy.

WL&P argued that the grammatical asymmetry is expected if comprehenders compute agreement using a cued-based retrieval mechanism. They proposed two alternative ways in which retrieval could be deployed during comprehension. One possibility is that retrieval functions as a repair or reanalysis mechanism triggered by the violation of a number prediction. On this view, the subject noun within the RC predicts the number of the verb. When the verb form violates this prediction, participants use cue-based retrieval to check whether the correct feature was somehow missed during first pass. Since the attractor "musicians" matches the verb in number, it is sometimes wrongly retrieved, which allows comprehenders to license of the verb and results in facilitated processing. In contrast, in the grammatical conditions the verb always matches the number prediction made by the subject noun, and

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

therefore retrieval is not engaged and no attraction takes place. Alternatively, comprehenders may always use cue-based retrieval to compute agreement when a verb is encountered, regardless of the grammaticality of the verb. Under this account, the grammatical asymmetry arises because in grammatical sentences, the RC subject "the reviewer" matches the verb in both number and syntactic cues, and thus it is retrieved most of the time, resulting in no interference from the attractor noun "musicians".

Importantly, both accounts involve cue-based retrieval and predict the existence of a grammatical asymmetry. This differentiates them from other accounts of agreement attraction, which predict symmetric attraction effects. For example, percolation accounts posit that attraction results from faulty movement or "percolation" of plural features from the attractor to the subject noun (Eberhard, Cutting, & Bock, 2005; Franck, Vigliocco, & Nicol, 2002; Pearlmutter et al., 1999). This causes the plural feature of the attractor to sometimes overwrite the number of the subject phrase. When this happens, speakers should wrongly represent the number of the subject phrase and symmetric attraction effects should occur: grammatical sentences with a plural attractor and a singular verb (2b) should sometimes be perceived as ungrammatical (yielding processing difficulty), while ungrammatical sentences with a plural attractor and a plural verb (2d) should sometimes be perceived as grammatical (yielding facilitation).

In this study, we adopt the view that agreement attraction in comprehension results from a retrieval mechanism and we use the grammatical asymmetry as a diagnostic to examine whether Spanish and English comprehenders process agreement errors in a similar way. In what follows, we consider three possible ways in which Spanish's richer morphology might affect the retrieval mechanism.

The present study

The goal of the present study was to determine whether a morphologically rich language like Spanish shows agreement attraction in comprehension. To our knowledge, while there have been several studies on agreement violations in morphologically rich languages (see Mancini, Molinaro, & Carreiras, 2013, for review), very few published comprehension studies have looked at subject–verb agreement attraction in languages other than English (Acuña-Fariña, Meseguer, & Carreiras, 2014). In contrast, there have been several cross-linguistic studies in production (Bock, Carreiras, & Meseguer, 2012; Franck, Vigliocco, Antón-Méndez, Collina, & Frauenfelder, 2008; Hartsuiker, Schriefers, Bock, & Kikstra, 2003; Lorimor, Bock, Zalkind, Sheyman, & Beard, 2008; Vigliocco, Butterworth, & Garrett, 1996; Vigliocco, Butterworth, & Semenza, 1995; Vigliocco, Hartsuiker, Jarema, & Kolk, 1996). However, the role of morphological richness in these studies remains unresolved, as some have found higher attraction rates in languages with poor morphology, whereas others have found the opposite pattern (see Lorimor et al., 2008, for discussion).

We examined three possible ways in which a richer morphology could affect the retrieval mechanism, causing Spanish and English speakers to behave differently. One possibility is that Spanish speakers could be less susceptible to attraction than English speakers. If retrieval is only engaged as a reanalysis strategy, it might be attempted less frequently in a language with richer morphology. In English, number morphology is used sparingly in nominal phrases. Therefore, English comprehenders might be less certain about their encoding of the previous input, making them more likely to attempt reanalysis and resulting in increased susceptibility to attraction. In Spanish, in contrast, plural number is clearly marked on both the determiner and attractor noun, and plural verb forms are different than bare forms. Therefore, Spanish comprehenders might be more certain about their encoding of the previous input such that they attempt reanalysis less often, resulting in smaller or non-existent attraction effects.

A second possibility that is related to theoretical accounts of cue usage (e.g. MacWhinney, 1987; MacWhinney & Bates, 1989) is that a richer verb morphology could make Spanish comprehenders more susceptible to attraction than English comprehenders. If richer morphology affects retrieval through differential cue weighting, and if morphological cues are weighted more heavily in Spanish, then comprehenders might make more errors at retrieval, resulting in larger attraction effects. This possibility is motivated by the fact that agreement cues are richer and more reliable in Spanish than in English, and therefore they might carry more weight at retrieval relative to syntactic cues, resulting in an increased number of incorrect retrievals of the attractor noun.

We contrast these two possibilities with the hypothesis that the retrieval mechanism is used uniformly across languages. If retrieval is implemented similarly cross-linguistically, regardless of differences in surface form, then similar attraction effects should be found in English and Spanish. Under this hypothesis, we expect Spanish speakers to show reduced processing cost for ungrammatical sentences that contain a plural attractor. Further, we expect to observe a grammatical asymmetry in Spanish, yielding facilitation in ungrammatical sentences but no processing difficulty in grammatical sentences. Overall, the presence of qualitatively and quantitatively similar attraction effects in Spanish and English would suggest that retrieval is implemented uniformly across the two languages.

In addition, we contrasted the two alternative accounts of retrieval that were outlined above, by comparing the temporal profiles of agreement violations and attraction effects. Under an account that posits that retrieval is an error-driven mechanism, comprehenders should use retrieval as a repair strategy only after they detect a subject–verb number mismatch. Thus, this account predicts that awareness of grammaticality violations should precede attraction effects. In contrast, an account that posits that retrieval is always engaged when a verb is encountered, assumes that comprehenders can only compute agreement after retrieval has taken place. Thus, this account predicts synchronic or simultaneous effects of grammaticality violations and attraction.

To address these alternatives, we used vincentile plots to compare the temporal onsets of agreement violations

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

and attraction effects. Two previous studies on agreement attraction using a forced-choice paradigm (Staub, 2009, 2010) found that attraction effects in relative clauses result from a process that involves only a small proportion of the trials, and it is seen in the right tail of a reaction time distribution. We wanted to examine whether a similar distributional pattern was obtained using a reading paradigm that did not require participants to make conscious choices between alternative verb forms.

Overview of the experiments

The experiments addressed whether Spanish comprehenders are susceptible to attraction effects, and whether their reading profiles are similar to English. To assess the effect of morphological richness on agreement attraction, we varied the markedness of the verb forms. Previous English experiments with relative clauses used main verbs with singular and plural forms differing in only one character (e.g. "praises vs. praise"). Experiment 1 used main verbs in Spanish that contrasted sharply in singular and plural forms. Experiments 2 and 3A and 3B used auxiliary verbs in Spanish and English, which are more closely matched in length between singular and plural forms. In all four studies we complement the mean reading time results with vincentile analyses of the reaction time distributions of grammaticality and attraction effects.

Experiment 1: Spanish main verbs

The goal of Experiment 1 was to investigate agreement attraction in Spanish comprehension. If the same retrieval mechanism causes attraction across languages, then English and Spanish comprehenders should show similar reading profiles, with plural attractors facilitating processing only in ungrammatical sentences, as previously shown in English by WL&P. Furthermore, if retrieval is deployed as a repair mechanism when comprehenders encounter a number violation, then grammatical violations should affect reaction time distributions earlier than attraction effects. In order to address this prediction, we performed distributional analyses using a non-parametric technique, vincentile plots, described below.

Participants

Participants ($n = 32$, mean age = 27 years, 15 females) were all native speakers of Argentinian Spanish and were recruited from the University of Buenos Aires community. All participants provided informed consent and were compensated the equivalent of $5/h.

Materials and design

Spanish materials were constructed based on the items in Experiment 2 of WL&P. They consisted of 48 sentence sets arranged in a 2×2 within-subjects design, with grammaticality (*grammatical*/*ungrammatical*) and attractor number (*singular*/*plural*) as factors. In the grammatical conditions, the subject and the relative clause verb were both singular (i.e., they agreed in number), while in the ungrammatical conditions the subject was singular and the verb was plural (i.e., they mismatched in number). The relative clause verbs were in the simple past tense and perfective aspect. The singular suffix for this tense-aspect combination in Spanish is one character long (e.g. "escribi-ó", 'write.3sg'), while the plural suffix is 4 characters long (e.g. "escribi-eron", 'write.3pl'). An adverbial phrase consisting of two or three words was introduced after the relative clause verb to allow for the presence of spillover effects (Just & Carpenter, 1978).

The attractor noun was the head noun modified by the relative clause. It was always inanimate (e.g. "nota", 'note') and either singular or plural. In contrast, the subject noun within the relative clause was always singular and animate (e.g. "chica", 'girl'). The 48 sentence sets were distributed across four lists in a Latin Square design, and were combined with 24 items (half ungrammatical) from a different experiment not reported here, and 188 grammatical filler sentences of similar length. This resulted in 13.8% of the items being ungrammatical. An example set of experimental materials is presented in Table 1 and the full item sets are available in Appendix A.

Procedure

Sentences were presented on a laptop PC using the Linger software (Doug Rohde, MIT) in a self-paced word-by-word moving window paradigm (Just, Carpenter, & Woolley, 1982). Each trial began with a screen presenting a sentence in which the words were masked by dashes. Each time the participant pressed the space bar, a word was revealed and the previous word was re-masked. The time spent on each word was measured as the time difference between two successive key presses. Every experimental and filler item was followed by a yes/no comprehension question to ensure that participants were attending to the stimuli. The comprehension questions never referred to the agreement dependency. On-screen feedback was provided for incorrect answers. Participants were instructed to read at a natural pace and answer the questions as quickly and accurately as possible. Three practice items were presented before the beginning of the experiment. Participants were not informed that sentences would contain grammatical errors. The order of experimental and filler items was randomized for each participant. An entire experimental session lasted approximately 1 h.

Analysis

Comprehension questions

A statistical analysis of the proportion of correct responses in the experimental trials was carried out using mixed effects logistic regression (Jaeger, 2008), with grammaticality, attractor number and their interaction as fixed effects. For consistency, the random effects structure of this model and the model used to analyze reaction times were identical (see below). Analyses were carried out using R, an open source programming language and environment for statistical computing (R Development Core Team,

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

Table 1
Sample set of experimental items in Experiments 1 and 3A.

Spanish main verbs	
Gram, sg attractor	**La nota** que la chica escribió en la clase alegró a su amiga *The note that the girl wrote$_{sg}$ during class cheered her friend up*
Gram, pl attractor	**Las notas** que la chica escribió en la clase alegraron a su amiga *The notes that the girl wrote$_{sg}$ during class cheered her friend up*
Ungram, sg attractor	*__La nota__ que la chica escribieron en la clase alegró a su amiga *The note that the girl wrote$_{pl}$ during class cheered her friend up*
Ungram, pl attractor	*__Las notas__ que la chica escribieron en la clase alegraron a su amiga *The notes that the girl wrote$_{pl}$ during class cheered her friend up*
Spanish auxiliary verbs	
Gram, sg attractor	**La nota** que la chica va a escribir en la clase alegrará a su amiga *The note that the girl is going to write during class will cheer her friend up*
Gram, pl attractor	**Las notas** que la chica va a escribir en la clase alegrarán a su amiga *The notes that the girl is going to write during class will cheer her friend up*
Ungram, sg attractor	*__La nota__ que la chica van a escribir en la clase alegrará a su amiga *The note that the girl are going to write during class will cheer her friend up*
Ungram, pl attractor	*__Las notas__ que la chica van a escribir en la clase alegrarán a su amiga *The notes that the girl are going to write during class will cheer her friend up*

2014), and in particular the *lme4* package for linear mixed effects models (Bates, Maechler, Bolker, & Walker, 2014). Only data from participants with accuracy above 70% in the filler items were included in the analysis.

Reaction times

To allow for comparison with the English study by WL&P our analysis was maximally similar to theirs. Reaction times (RTs) that exceeded a threshold of 2.5 standard deviations by region and condition were excluded (Ratcliff, 1993). Across all experiments reported in this paper, on average this resulted in the exclusion of 2.4% of trials in the critical regions (range = 2.2–3.1%). Based on previous studies, we identified two regions of interest: the verb in the relative clause (*verb region*) and the word immediately following it (*verb+1 region*).

RTs were analyzed using a model with grammaticality, attractor number and their interaction as fixed effects. Both main effects were coded using orthogonal contrasts: for the grammaticality factor, the mean of the ungrammatical conditions was compared with the mean of the grammatical conditions. For the attractor number factor, the mean of the plural noun conditions was compared with the mean of the singular noun conditions. Following current guidelines in the psycholinguistics literature (Barr, Levy, Scheepers, & Tily, 2013), we initially constructed a maximal model that included random intercepts and slopes for all fixed effects and their interaction. However, as this maximal model failed to converge in most cases, we gradually simplified the random effects structure following the suggestions by Barr and colleagues. We report the results from the model with the maximal random effects structure that converged for all critical regions across experiments and that did not contain correlations between the random effects equal to 1 (Baayen, Davidson, & Bates, 2008). The final model included by-subject and by-item random intercepts. The tables show the model estimates in milliseconds ($\hat{\beta}$), with negative values corresponding to a decrease in RTs. *p*-Values were computed using Satterthwaite's approximation for denominator degrees of freedom with the *lmerTest* package (Kuznetsova, Bruun Brockhoff, & Haubo Bojesen Christensen, 2014). In order to ensure consistency with WL&P, we also performed a complementary analysis computing ANOVAs by subjects (F_1) and items (F_2). This analysis yielded similar results to the linear mixed effects model and is presented in Appendix B.

Finally, one difference between WL&P's experiment and the current study is that due to Spanish verbal morphology, ungrammatical verbs were on average three characters longer than grammatical verbs. This is problematic, as both ungrammaticality and word length have been shown to increase reaction times. Therefore, in order to avoid a confound between these two factors, the effect of word length was estimated from the entire dataset, and then regressed from the raw reaction times using a linear model (cf. Hofmeister, 2010). Note that this correction may have been over-conservative since additional character length in verbal suffixes, which are closed class morphemes, might not increase reading times to the same degree as additional length does in root morphemes. For instance, character length has been found to have a larger effect on low than high frequency words in some previous lexical processing studies (e.g. Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Assuming that plural/singular suffixes are more frequent than root morphemes, the length correction may be over-conservative and might have made it harder to find grammaticality effects. However, we chose this method to avoid any confound between length and grammaticality. Only the length-regressed reaction times were entered into the statistical analysis.

Vincentile plots

Vincentile plots were used to examine the effect of each experimental factor on participants' reaction time distributions (Ratcliff, 1979; Vincent, 1912). Vincentile plots

are a non-parametric way of assessing the shape of a reaction time distribution. They are plots of quantiles, estimated in a way that is robust to outliers and supports averaging across participants. They were constructed as follows. First, as the entire RT distribution was of interest, a conservative trimming procedure was used, excluding only RTs that exceeded 4000 ms (cf. Staub, 2010). This resulted in the exclusion of 0.1% of the trials in the critical regions (range = 0–0.4%) across the experiments reported in this paper.

The raw RTs for each participant in each condition were binned into the shortest 10% (vincentile 1), the next shortest 10% (vincentile 2), etc. The mean of the observations in each vincentile was calculated. Finally, vincentile plots were computed by collapsing the same bins across participants. Note that since RTs were assigned to vincentiles on a per-participant basis, data from faster and slower readers are evenly represented across vincentiles. These values were displayed as connected points on a plot with vincentile number on the *x*-axis and reaction time in milliseconds on the *y*-axis.[1]

In order to compare grammaticality and attraction effects we computed three difference scores: a *grammaticality difference* (computed by subtracting the grammatical from the ungrammatical vincentile curve in the no attractor conditions), an *ungrammatical attractor difference* (obtained by subtracting the plural from the singular vincentile curve in the ungrammatical conditions) and a *grammatical attractor difference* (obtained by subtracting the plural from the singular vincentile curve in the grammatical conditions). Therefore, each of these three difference scores consisted of ten values per participant and they expressed the effects of grammaticality, attraction in the ungrammatical conditions and attraction in the grammatical conditions respectively.

The difference scores were analyzed with a linear model with vincentile number as a categorical predictor and with grammaticality (for the *grammaticality difference*) and attractor number (in the *attractor ungrammatical* and *grammatical attractor differences*) as fixed effects and by-subject random intercepts. For each contrast, we compared the difference score in each vincentile with the mean difference score in the previous vincentiles (e.g., the difference score in the second vincentile was compared with the difference score in the first vincentile, the difference score in the third vincentile was compared with the mean difference score of the first and second vincentiles, etc.). We defined the *divergence point* as the earliest vincentile with a difference score significantly different from preceding vincentiles. As all vincentiles following the divergence point also displayed significant differences, we only report the statistics associated with the earliest vincentile where the difference became significant. Note that vincentiles were not compared across conditions, but rather, each vincentile was compared against previous vincentiles in the same condition. This was done to avoid biasing ourselves to find earlier onsets for effects of larger size overall: since grammaticality effects were always larger than attraction effects in mean reaction times, comparing vincentiles across conditions would have introduced a bias for finding an earlier onset for the largest effect. The comparison of each vincentile with preceding vincentiles within the same condition was designed to provide an appropriate baseline to each vincentile difference in the statistical analysis.

It is important to clarify that in previous sentence comprehension studies (Staub, 2011; Staub, White, Drieghe, Hollway, & Rayner, 2010) vincentiles have been often used to complement parametric analyses such as the fitting of an ex-Gaussian distribution to reading time data. In this study, however, due to the restricted number of trials per participant per condition ($n = 12$), we could not conduct parametric analysis of the reaction time profiles. In contrast with parametric techniques, the use of vincentiles does not involve making assumptions about the underlying distribution of the observed reaction time data. Therefore, the vincentile plots are provided here with the goal of (i) illustrating the impact of grammaticality and attraction effects on the shape of the reaction time distributions, and (ii) identifying the earliest point where each variable impacted participants' reaction times.

Results

In the comprehension questions, mean accuracy was 95.1% and no participants were excluded due to low accuracy. Question accuracy did not differ by condition.

The region-by-region reaction time averages are presented in Fig. 1. We plot raw RTs for easier readability, but the statistical analysis was always performed on length-regressed RTs. Plots for length-regressed RTs for this and following experiments can be found in Appendix C. The results from the mixed-effects model in the verb and the verb+1 regions are shown in Table 2. At the verb there was a marginal main effect of grammaticality, with the ungrammatical conditions having longer RTs than the grammatical conditions. This marginal effect contrasts with the large slowdown seen in the plot. However, the ungrammatical verbs were also markedly longer than the grammatical verbs. Since the effect of grammaticality was marginal after length was regressed from the RTs, this suggests that the difference in the plot was partly due to the length of the verbs.

At the verb+1 region, there was a main effect of grammaticality, a main effect of attractor number and a significant interaction between them. The ungrammatical conditions had longer RTs than the grammatical conditions, and the plural attractor conditions had shorter RTs than the singular noun conditions. Crucially, the interaction shows that the number of the attractor noun affected grammatical and ungrammatical sentences differently. Pairwise comparisons indicated that in ungrammatical sentences the plural attractor condition was read significantly faster than the condition with a singular noun

[1] One possible concern is that due to the number of trials per participant per condition in the experiment ($n = 12$) the use of 10 vincentiles means having on average one data point per participant per bin, which is less than the number of observations traditionally used in vincentile analyses. To address this concern, we conducted a complementary analysis using only 4 vincentiles, which allowed having 3 observations per participant per bin. As the two analyses yielded very similar results, we present the analysis based on 10 vincentiles and we note any discrepancies in the text.

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

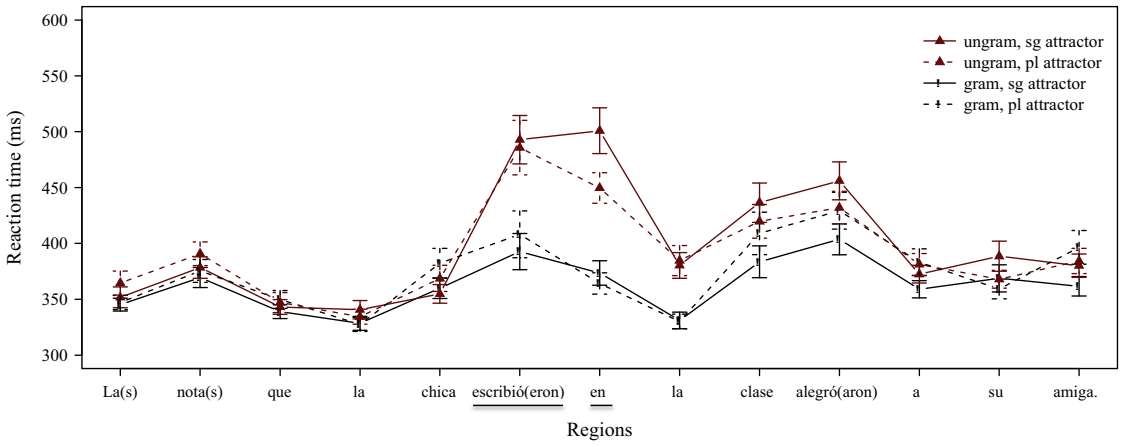

Fig. 1. Region-by-region means in Experiment 1. Error bars indicate the standard error of the mean. Sample sentence: *The note(s) that the girl wrote*$_{sg/pl}$ *during class cheered her friend up*. The regions of interest are underlined.

Table 2
Linear mixed-effects model results for the regions of interest in Experiment 1.

	$\hat{\beta}$	*SE*	*t*	*p*
Verb region				
Grammaticality	***18***	***10***	***1.80***	***<.07^***
Attractor number	−4	10	−0.40	.69
Gram × Attr number	−15	20	−0.77	.44
Verb+1 region				
Grammaticality	**93**	**8**	**11.99**	**<.01****
Attractor number	**−23**	**8**	**−2.94**	**<.01****
Gram × Attr number	**−34**	**15**	**−2.20**	**<.05***

($\hat{\beta} = -39$, $t = -2.93$, $p < .01$). In contrast, in grammatical sentences the number of the attractor did not significantly affect RTs ($\hat{\beta} = -6$, $t = -0.80$, $p = .42$).

In the vincentile analysis, we examined the reaction time distributions in the post-verbal region, where both grammaticality and attraction effects were observed. The vincentiles show how a reaction time distribution changes across conditions. For example, if a difference between two condition means is mainly due to a change that affects most trials, then the vincentiles corresponding to two different conditions should be parallel to each other across all vincentiles, with the curve corresponding to the slower condition appearing above the curve representing the faster condition. If, on the other hand, a difference between two conditions is due to a change that affects only the slower trials (i.e., the trials with the longest RTs), then the vertical distance between the vincentiles should be small or nonexistent for the vincentiles on the left-hand side of the graph, but larger on the right. Finally, a difference that is due to a change that affects most trials, but has a disproportionate effect on the slower trials, should show a vertical separation that is present across the full range of vincentiles but that is larger on the right than on the left.

Fig. 2 shows the vincentile plot and the results from the statistical analyses are reported in the text. Visual inspection of the plot suggested that grammaticality and

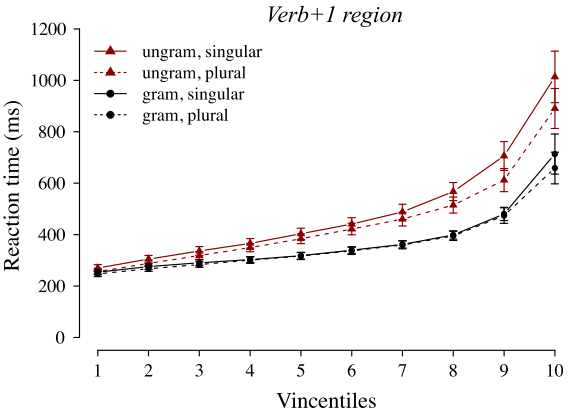

Fig. 2. Vincentile plots for the four experimental conditions in the verb+1 region in Experiment 1. Error bars indicate the standard error of the mean across participants.

attraction effects affected the late vincentiles more strongly, consistent with a process confined to the subset of slower RTs. This pattern was tested using an ANOVA with vincentile number as a factor (cf. Balota, Yap, Cortese, & Watson, 2008). The results confirmed the visual pattern, as shown by a vincentile number × grammaticality interaction for the *grammaticality difference* ($\hat{\beta} = 29$, $t = 5.78$, $p < .01$), and a vincentile number × attractor number interaction in the *ungrammatical attractor difference* ($\hat{\beta} = -11$, $t = -2.04$, $p < .05$). The grammatical conditions, on the other hand, were closely aligned and no interaction was found, consistent with the lack of a significant effect in the overall means.

Secondly, we examined the onsets of grammaticality and attraction effects by comparing the difference score in each vincentile with the difference score in the preceding vincentiles. The divergence point was defined as the earliest vincentile where the difference from the preceding vincentiles in the same condition was significant. Grammatical and ungrammatical vincentile curves differed from the seventh vincentile onwards (*grammaticality*

difference: $\hat{\beta}_{V7} = 70$, $t = 2.04$, $p < .05$). As predicted, the attraction effect in the ungrammatical conditions had a later onset, being marginal in the ninth vincentile and significant in the last vincentile (*ungrammatical attractor difference:* $\hat{\beta}_{V10} = 92$, $t = 2.39$, $p < .05$). In the grammatical conditions, there was also an effect in the last vincentile (*grammatical attractor difference:* $\hat{\beta}_{V10} = 50$, $t = 2.23$, $p < .05$).[2]

Discussion

This experiment shows clear attraction effects in Spanish comprehension. Consistent with the predictions of a retrieval account, we replicated the grammatical asymmetry in Spanish and found that plural nouns only facilitated processing in ungrammatical sentences, as evidenced by the attractor number × grammaticality interaction. This suggests the processing of ungrammatical sentences was eased when there was a syntactically unlicensed but number matching attractor in the sentence.

The vincentiles enriched the conclusions from the analysis of the mean RTs. First, they showed that the attraction effect impacted late vincentiles more strongly, as evidenced by the vincentile number × attractor number interaction in the ungrammatical conditions. This result supports previous findings (Staub, 2009, 2010) that have argued that attraction effects in RCs are driven by a small set of trials, namely, those that have disproportionately longer reaction times.

In addition, grammaticality affected the verb+1 vincentiles earlier than the attraction effect in ungrammatical sentences. This was paralleled in the standard region-by-region reading time measures, where the grammaticality effect was marginally significant at the verb while the attraction effect in ungrammatical sentences was not visible until the verb+1 region. The earlier effect of grammaticality relative to attraction is consistent with the hypothesis that the cognitive processes associated with the detection of an agreement violation preceded those associated to attraction. This supports an error-driven retrieval mechanism in which comprehenders first detect an agreement violation and then probe the contents of their memory to find a plural-marked noun to license the verb. This timing profile is not predicted under the view that retrieval is always engaged when a verb is encountered; in this case, since the subject noun does not predict the verb number, agreement can only be computed after retrieval has taken place. However, we note that the change in the empirical reaction time distributions could also have arisen if both underlying processes had similar time-courses but affected the vincentiles with different strengths (see General discussion), and therefore the vincentile results should only be taken suggestive.

Overall, these results suggest that agreement attraction in Spanish results from a retrieval mechanism. If retrieval is used as a reanalysis process, these results suggest that comprehenders engage in retrieval even when clear and unambiguous morphological cues are present in the input. However, we also wanted to examine comprehension profiles using Spanish verbs with a weaker morphological marking, to allow a closer comparison between English and Spanish. Therefore our next studies used auxiliary verbs, since in Spanish the distinction between plural and singular auxiliary verbs is less morphologically salient and more closely matched to the previous English experiments.

[2] This effect was unexpected, as no difference was observed in the grammatical conditions in the mean RTs analysis. However, in contrast with the other comparisons, this difference was not replicated in the complementary analysis using only 4 vincentiles, so it may have been driven by a few outlier points in the 10-vincentile analysis.

Experiment 2: English auxiliaries

Our goal in the next studies was to examine attraction effects elicited by verbs with weaker morphological markings. This was achieved by replacing the main verbs inside RCs in Experiment 1 with auxiliaries that only differed in one character between singular and plural forms. However, before conducting the Spanish experiments, we wanted to ensure that any differences between experiments would be due to a contrast in morphological marking, and not to lexico-semantic differences between auxiliary and main verbs. Since previous English experiments never used auxiliaries in RC constructions, we first tested English auxiliaries to ensure that they elicited the same attraction profiles as main verbs.

Participants

Participants ($n = 32$, mean age = 21 years, 22 females) were all native speakers of English and were recruited from the University of Maryland community. All participants provided informed consent and were compensated $10/h or received course credit for their participation.

Materials and design

We created 48 experimental items by taking WL&P's sentences (Experiment 2), and inflecting the RC verbs in the past progressive forms (e.g. the main verb "praise" was replaced by "was/were praising"). To ensure similarity with the Spanish auxiliaries experiment, where a preposition appeared between the auxiliary and the main verb, we inserted an adverb between the auxiliary and the main verb (e.g. "was always praising"). WL&P's items were only modified when the main verb could not be used in the past progressive form (e.g. "know"). In these cases, the verb and its spillover region were changed, resulting in partial modification of half of the item sets. An example set of the experimental materials is presented in Table 3.

Procedure and analysis

The same self-paced reading procedure and analysis as in Experiment 1 were used. However, as there was one extra word between the auxiliary and the main verb, three regions of interest were defined: the auxiliary verb in the relative clause (*verb region*), the adverb following it (*verb+1 region*) and the non-finite main verb (*verb+2 region*).

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

Table 3
Sample set of experimental items in Experiment 2.

English auxiliary verbs	
Gram, sg attractor	**The player** that the coach was always praising very enthusiastically decided to leave the team
Gram, pl attractor	**The players** that the coach was always praising very enthusiastically decided to leave the team
Ungram, sg attractor	*__The player__ that the coach were always praising very enthusiastically decided to leave the team
Ungram, pl attractor	*__The players__ that the coach were always praising very enthusiastically decided to leave the team

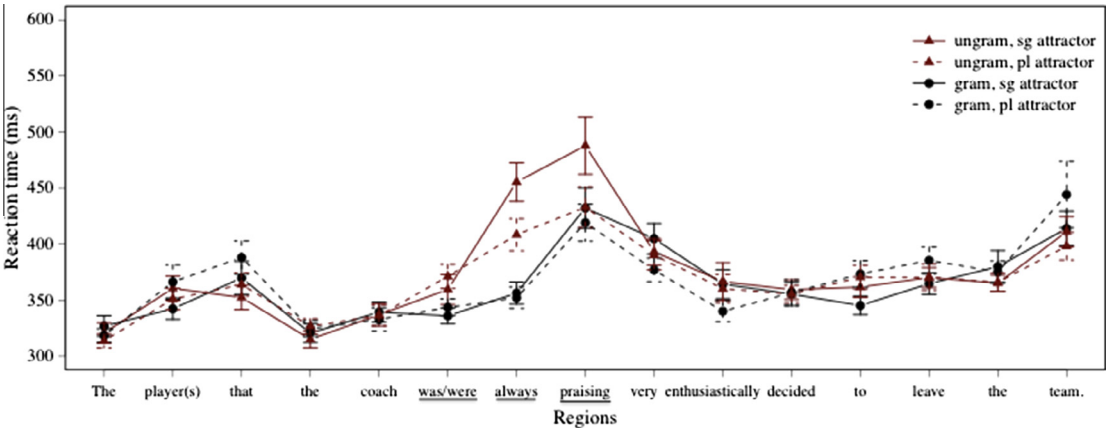

Fig. 3. Region-by-region means in Experiment 2. Error bars indicate the standard error of the mean. The regions of interest are underlined.

Results

In the comprehension questions mean accuracy was 94.1% and no participants were excluded due to low accuracy. Question accuracy did not differ across conditions.

The region-by-region reaction time averages are presented in Fig. 3. Length-regressed RTs were analyzed in the three regions of interest and the results from the mixed-effects model are shown in Table 4. At the auxiliary there was a main effect of attractor number, with the plural attractor conditions being read more slowly than the singular attractor conditions. The verb+1 region showed main effects of grammaticality, attractor number and a significant interaction between them. As predicted, ungrammatical conditions were read more slowly than grammatical conditions, and sentences with a plural attractor noun were read more quickly than sentences with a singular noun. Crucially, the interaction showed that plural attractors decreased RTs only in the ungrammatical conditions ($\hat{\beta} = -36$, $t = -2.94$, $p < .01$). In the grammatical sentences, there was no significant effect of attractor number ($\hat{\beta} = -4$, $t = -0.48$, $p = .63$). Finally, the two main effects continued in the verb+2 region, where ungrammatical conditions were read more slowly than grammatical conditions, and plural attractor conditions were read more quickly than singular noun conditions.

A vincentile plot was calculated for the verb+1 region, where grammaticality and attraction effects were observed (Fig. 4). The inspection of the plot revealed that grammaticality and attraction effects were more pronounced in the late vincentiles, as confirmed by a vincentile number × grammaticality interaction for the *grammaticality difference* ($\hat{\beta} = 29$, $t = 5.99$, $p < .01$), and a vincentile number × attractor number interaction for the *ungrammatical attractor difference* ($\hat{\beta} = -15$, $t = -2.81$, $p < .01$). In the grammatical sentences, no interaction was found.

Table 4
Linear mixed-effects model results for the regions of interest in Experiment 2.

	$\hat{\beta}$	*SE*	*t*	*p*
Verb region				
Grammaticality	2	5	0.35	.73
Attractor number	**12**	**5**	**2.27**	**<.05***
Gram × Attr number	9	10	0.84	.40
Verb+1 region				
Grammaticality	**49**	**7**	**6.66**	**<.01****
Attractor number	**−20**	**7**	**−2.71**	**<.01****
Gram × Attr number	**−32**	**15**	**−2.19**	**<.05***
Verb+2 region				
Grammaticality	**23**	**9**	**2.56**	**<.05***
Attractor number	**−23**	**9**	**−2.48**	**<.05***
Gram × Attr number	−31	18	−1.68	.09

Lastly, the grammaticality effect showed an earlier onset than the attraction effect. Ungrammatical and grammatical vincentile curves differed significantly from the seventh vincentile onwards (*grammaticality difference*: $\hat{\beta}_{V7} = 86$, $t = 2.92$, $p < .01$), whereas the attractor vs. no-attractor vincentile curves in the ungrammatical conditions differed from the eight vincentile onwards (*ungrammatical attractor difference:* $\hat{\beta}_{V8} = 69$, $t = 2.23$, $p < .05$). The *grammatical attractor difference* was not significant in any vincentile.

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

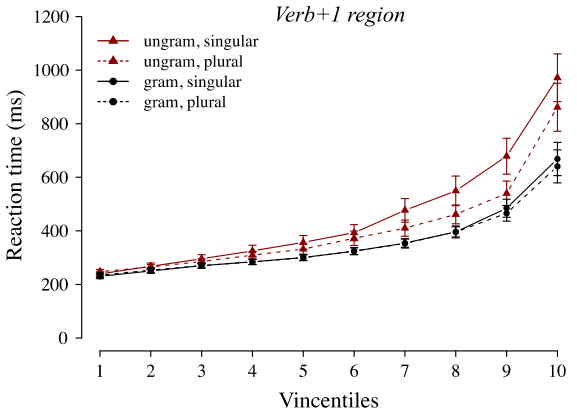

Fig. 4. Vincentile plots for the four experimental conditions in verb+1 region in Experiment 2. Error bars indicate the standard error of the mean across participants.

Discussion

These results show that English auxiliaries in relative clauses behave similarly to main verbs, which were previously tested by WL&P. Our findings showed a significant attraction effect in ungrammatical sentences, where processing was eased by the presence of a plural attractor. Importantly, no difference was observed in the grammatical sentences, as shown by the grammaticality × attractor number interaction. In addition, the vincentile plots revealed that the attraction effect impacted the late vincentiles more strongly and showed a later onset than the grammaticality effect. As in Experiment 1, these results are consistent with the hypothesis that English comprehenders employ error-driven retrieval when they encounter a number-marked auxiliary, as was argued for main verbs. The following experiment examined Spanish auxiliaries.

Experiment 3: Spanish auxiliaries

Experiment 3A

This experiment examined the contrast between main verbs and auxiliaries in Spanish, using materials identical to Experiment 1 but changing the nature of the verb within the relative clause.

Participants

Participants ($n = 32$, mean age = 27 years, 24 females) were all native speakers of Argentinian Spanish and were recruited from the University of Buenos Aires community. All participants provided informed consent and were compensated the equivalent of $5/h.

Materials and design

The design was identical to the previous experiments. The same sentence items as in the main verbs experiment were used, with the exception that main verbs were replaced by auxiliary constructions, which consisted of an auxiliary in the future tense followed by the preposition "a", followed by a non-finite main verb (e.g. "va(n) a escribir", 'is/are going to write'). An example set of the experimental materials is presented in Table 1.

Procedure and analysis

The procedure and analysis were identical to previous experiments with one exception. Due to the introduction of the future auxiliary construction "va a", there was one extra region for analysis: the regions of interest consisted of the auxiliary verb in the relative clause (*verb region*), the preposition "a" (*verb+1 region*) and the non-finite main verb following it (*verb+2 region*). Also, one participant who had less than 10 trials in one of the experimental conditions was excluded from the vincentile analysis.

Results

Mean accuracy in the comprehension questions was 95.1% and no participants were excluded due to low accuracy. Question accuracy did not differ across conditions.

Fig. 5 shows region-by-region reaction time averages and the mixed-effects model results are presented in Table 5. At the auxiliary there were no main effects or interactions. At the verb+1 region, there was a main effect of grammaticality and an interaction between grammaticality and attractor number. Pairwise comparisons showed that this interaction was due to the number of the attractor having opposite effects in the grammatical and ungrammatical conditions. In ungrammatical sentences the plural attractor condition elicited shorter RTs than the singular noun condition ($\hat{\beta} = -15$, $t = -2.23$, $p < .05$). In contrast, in grammatical sentences the plural attractor condition was associated with longer RTs than the singular noun condition ($\hat{\beta} = 12$, $t = 2.80$, $p < .01$). In the following verb+2 region, only a main effect of grammaticality was observed.

A vincentile plot was calculated for the verb+1 region, where both grammaticality and attraction effects were observed (Fig. 6). The vincentile curves displayed patterns that were consistent with the mean RT analysis, and opposite effects of attraction were observed in grammatical and ungrammatical sentences. Most of the effects seemed to impact the late vincentiles more strongly, as confirmed by a vincentile number × grammaticality interaction for the *grammaticality difference* ($\hat{\beta} = 12$, $t = 3.60$, $p < .01$), and a vincentile number × attractor number interaction for the *grammatical attractor difference* ($\hat{\beta} = -6$, $t = -2.17$, $p < .05$). In the ungrammatical sentences, however, the interaction did not reach significance. Inspection of the plot suggested that the plural attractor condition was still associated with a downward deflection as compared with the singular condition. Although this deflection was numerically largest in the last vincentile, the difference was smaller than in Experiment 1 and appeared to be more evenly distributed across vincentiles. However, it is possible that the overall smaller effect size of attraction in this experiment reduced our ability to detect an interaction.

Finally, the grammaticality effect showed an earlier onset than the attraction effect in the grammatical conditions. For the *grammaticality difference*, the grammatical and ungrammatical vincentile curves differed from the eighth vincentile onwards ($\hat{\beta}_{V8} = 45$, $t = 2.31$, $p < .05$). For

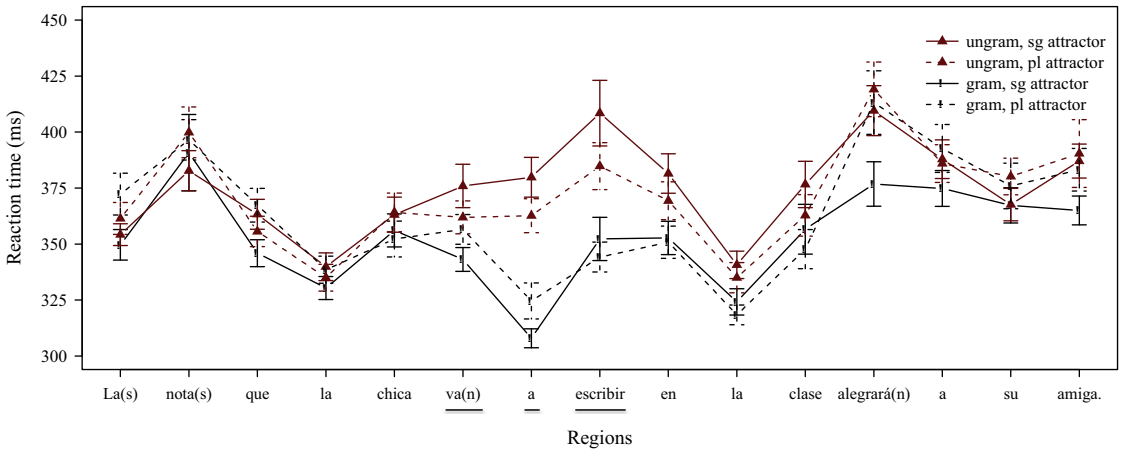

Fig. 5. Region by region means in Experiment 3A. Error bars indicate the standard error of the mean. Sample sentence: *The note(s) that the girl is/are going to write during class will cheer her friend up*. The regions of interest are underlined.

Table 5
Linear mixed-effects model results for the regions of interest in Experiments 3A and 3B.

	$\hat{\beta}$	*SE*	*t*	*p*
Experiment 3A				
Verb region				
Grammaticality	0	4	0.10	.93
Attractor number	3	4	0.70	.48
Gram × Attr number	−11	8	−1.35	.18
Verb+1 region				
Grammaticality	**47**	**4**	**11.36**	**<.01****
Attractor number	−2	4	−0.51	.61
Gram × Attr number	**−27**	**8**	**−3.25**	**<.01****
Verb+2 region				
Grammaticality	**39**	**5**	**7.12**	**<.01****
Attractor number	−7	5	−1.25	.21
Gram × Attr number	−4	11	−0.34	.73
Experiment 3B				
Verb region				
Grammaticality	**17**	**7**	**2.5**	**<.05***
Attractor number	−9	7	−1.4	.2
Gram × Attr number	5	13	0.4	.7
Verb+1 region				
Grammaticality	**56**	**5**	**10.86**	**<.01****
Attractor number	**−14**	**5**	**−2.80**	**<.01****
Gram × Attr number	−14	10	−1.36	.18
Verb+2 region				
Grammaticality	13	7	1.81	.07
Attractor number	8	7	1.08	.28
Gram × Attr number	24	15	1.60	.11

the *grammatical attractor difference*, the singular and plural noun vincentiles only differed in the last vincentile ($\hat{\beta}_{V10} = -95$, $t = -4.78$, $p < .01$). For the *ungrammatical attractor difference*, there were no significant differences.

Discussion

Our results in the ungrammatical conditions replicated the pattern seen for Spanish main verbs: comprehenders showed facilitated processing of ungrammatical sentences when they contained a syntactically unlicensed but number-matching noun. The vincentiles showed that plural attractors had opposite effects in grammatical and ungrammatical sentences. In grammatical sentences, attraction increased the slope of the plural attractor curve in the late vincentiles. In ungrammatical sentences, attraction decreased the slope of the plural attractor curve. Somewhat unexpectedly, in this case we observed no interaction between vincentile and attractor number, although it is possible that the notably smaller attraction effect in this experiment relative to Experiment 1 limited our ability to detect an interaction.

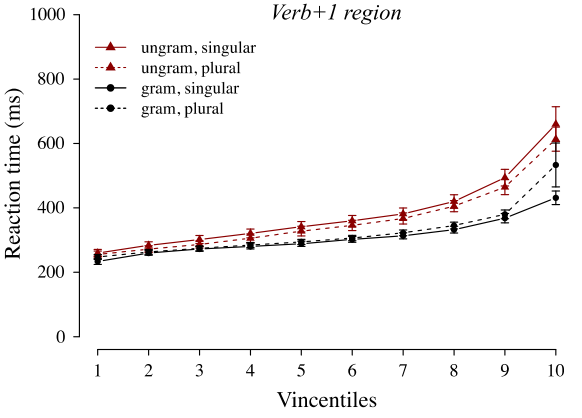

Fig. 6. Vincentile plots for the four experimental conditions in the verb+1 region in Experiment 3A. Error bars indicate the standard error of the mean across participants.

Taken together with the main verb findings, these results suggest that Spanish comprehenders use retrieval uniformly despite differences in morphological markedness within a language. In ungrammatical sentences, plural attractors eased processing after the verb, regardless of whether ungrammatical verbs had stronger (Experiment 1) or weaker morphological markings (Experiment 3A). This suggests that both within and across-languages, cue-based retrieval is used to repair grammatical violations, resulting in attraction from number-matching attractor nouns.

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

The unexpected finding in this experiment concerns the grammatical conditions, where comprehenders showed evidence of processing cost in sentences with plural attractors. This pattern is inconsistent with an error-driven retrieval account, under which attraction arises due to a retrieval mechanism triggered by the detection of an agreement mismatch. As RC subjects and verbs always agreed in number in the grammatical sentences, retrieval should not have been engaged and no attraction should have been observed. In contrast, attraction in grammatical sentences is expected under percolation accounts. Under a percolation account, comprehenders sometimes incorrectly encode the number of the subject phrase when it contains a plural attractor. Therefore, when they read the RC verb, they should sometimes mistakenly perceive ungrammatical sentences as grammatical (yielding facilitated processing), and grammatical sentences as ungrammatical (yielding processing cost).

Note, however, that a percolation view cannot account for the results of the main verb and auxiliary experiments taken together. In Experiment 1, where main verbs were used instead of auxiliaries, grammatical sentences did not show any evidence of processing cost. Since percolation accounts assume that feature-spreading takes place during processing of the subject phrase (before a verb is encountered) they cannot predict a dissociation contingent on the type of verb. In other words, since main verbs and auxiliaries had identical sentence preambles in Experiments 1 and 3A, similar results should have been obtained under a percolation account.

Alternatively, the attraction effect in the grammatical conditions may have been due to a Type I error. This concern is motivated by the fact that this pattern has not typically been observed in English (cf. Wagers et al., 2009, for discussion) and it was absent for main verbs in Spanish. Furthermore, the effect was found in a region that corresponded to a very short word (the particle "a" in Spanish), and was no longer significant when we conducted a supplementary analysis merging this region with the following region. In order to examine the reliability of this effect we conducted a supplementary experiment.

Experiment 3B

This experiment assessed the reliability of the attraction effect in grammatical sentences in Experiment 3A. Additionally, we tried to maximize the precision of the across-language comparison, and thus we translated the English sentences from Experiment 2, to ensure that sentences in Spanish and English were lexically matched.

Participants

Participants ($n = 32$, mean age = 24 years, 21 females) were all native speakers of Argentinian Spanish and were recruited from the University of Mendoza community in central Argentina. All participants provided informed consent and were compensated the equivalent of $5/h.

Materials and design

The design was identical to the previous experiments. The experimental sentences were translations from the English items in Experiment 2. Sentences were only changed when the English versions seemed implausible or unnatural in Spanish (e.g. "quarterback" was replaced by "swimmer"). This resulted in the modification of 8 nouns in the RC constructions. As in Experiment 3A, the RC verbs consisted of future auxiliary constructions (e.g. "va(n) a felicitar", 'is/are going to praise'). The materials are available in Appendix A.

Procedure and analysis

The procedure and analysis were identical to previous experiments. The regions of interest consisted of the auxiliary verb in the relative clause (*verb region*), the preposition "a" (*verb+1 region*) and the non-finite main verb following the preposition (*verb+2 region*).

Results

Mean accuracy in the comprehension questions was 94% and no participants were excluded due to low accuracy. Question accuracy did not differ across conditions.

Fig. 7 shows region-by-region reaction time averages and the mixed-effects model results are presented in Table 5. The auxiliary verb showed a main effect of grammaticality, with longer reaction times in the ungrammatical than in the grammatical conditions. At the verb+1 region, there were main effects of grammaticality and attractor number. The effect of attractor number was due to longer reaction times in the plural noun than in the singular noun conditions. The interaction between grammaticality and attractor number did not reach significance, but we conducted pairwise comparisons in the grammatical and ungrammatical conditions separately, as it was driven by our hypothesis of interest and by the visual pattern in the plot. These comparisons revealed an attraction effect in the ungrammatical conditions, where the plural attractor condition elicited shorter RTs than the singular noun condition ($\hat{\beta} = -21$, $t = -2.40$, $p < .05$). However, there was no significant difference in the grammatical conditions ($\hat{\beta} = -7$, $t = -1.61$, $p = .11$), and the pattern of means (plural attractor > singular attractor) was the opposite of what would be predicted for grammatical attraction. No effects were observed in the verb+2 region.

A vincentile plot was calculated for the verb+1 region, where both grammaticality and attraction effects were observed (Fig. 8). All experimental effects impacted the late vincentiles more strongly, as confirmed by a vincentile number × grammaticality interaction for the *grammaticality difference* ($\hat{\beta} = 28$, $t = 6.61$, $p < .01$) and vincentile number × attractor number interactions for the un*grammatical attractor difference* ($\hat{\beta} = -12$, $t = -2.57$, $p < .05$) and the *grammatical attractor difference* ($\hat{\beta} = 6$, $t = 3.1$, $p < .01$).

The grammaticality effect showed an earlier onset than the attraction effect. For the *grammaticality difference*, the difference between grammatical and ungrammatical vincentile curves was marginal at vincentile eight and significant from the ninth vincentile onwards ($\hat{\beta}_{V9} = 91$, $t = 3.35$, $p < .01$). In contrast, the singular and plural noun conditions only differed in the last vincentile for the *ungrammatical attractor* ($\hat{\beta}_{V10} = 156$, $t = 4.77$, $p < .01$) and the

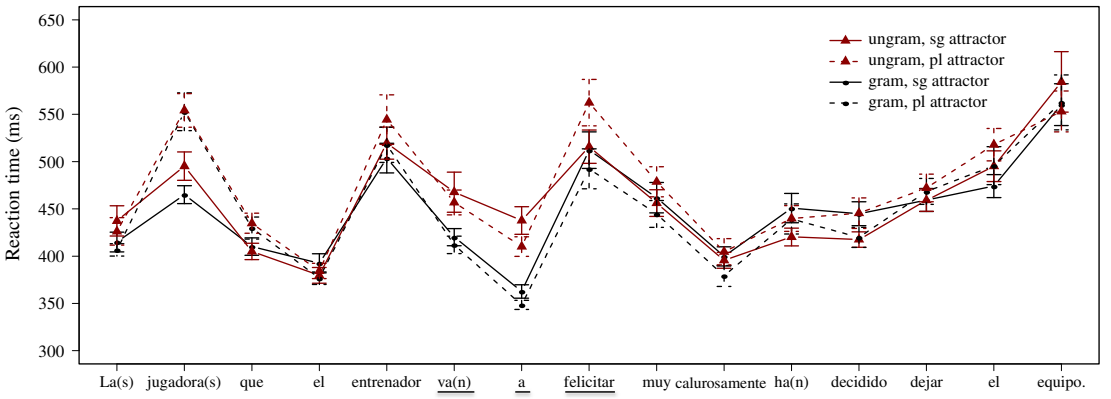

Fig. 7. Region by region means in Experiment 3B. Error bars indicate the standard error of the mean. Sample sentence: *The player(s) that the coach was/were always praising very enthusiastically has/have decided to leave the team.* The regions of interest are underlined.

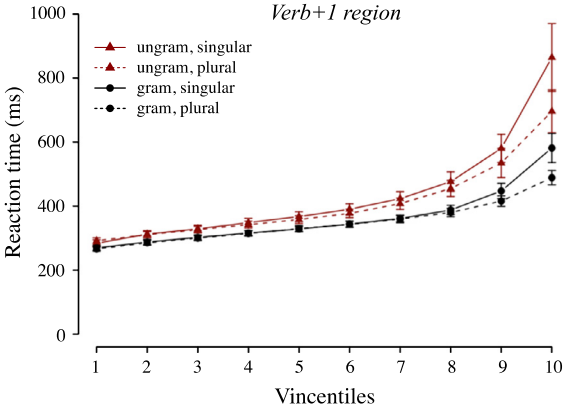

Fig. 8. Vincentile plot for the four experimental conditions in the verb+1 region in Experiment 3B. Error bars indicate the standard error of the mean across participants.

grammatical attractor differences ($\hat{\beta}_{V10} = 87$, $t = 5.49$, $p < .01$). In contrast with Experiment 3A, however, both grammatical and ungrammatical conditions showed the same pattern, with attraction decreasing the slope of the plural attractor curve.

Discussion

Experiment 3B aimed to address the reliability of attraction effects in grammatical sentences with Spanish auxiliary verbs, and to conduct a closer comparison between Spanish and English by using materials that were lexically matched across languages. The mean reaction time analysis revealed an attraction effect after the auxiliary verb in ungrammatical sentences, but there was no evidence of attraction in grammatical sentences. In fact, the plural noun condition showed shorter reading times than the singular noun condition in the two regions after the auxiliary verb. This pattern is inconsistent with the results of Experiment 3A, and with the predictions of a percolation account, under which plural attractors in grammatical sentences should elicit processing difficulty. In Experiment 3B, the trend in the *verb+1* and *verb+2* regions was the reverse, which is more consistent with processing facilitation than with processing difficulty.

As in previous experiments, the vincentiles revealed that attraction effects impacted the late vincentiles more strongly and showed a later onset than grammaticality effects. Unexpectedly, in this experiment there was also a difference in the vincentiles of the grammatical conditions. This difference was similar to the difference in the ungrammatical vincentiles, and it differed from Experiment 3A, as the plural noun vincentile showed a decreased slope as compared with the singular noun vincentile.

Overall, our results suggest that attraction effects in ungrammatical sentences are robust in Spanish, as they were observed across both Experiments 3A and 3B. In contrast, attraction effects in grammatical sentences were more unreliable: the processing difficulty associated with plural attractors in Experiment 3A was not replicated in Experiment 3B, and furthermore, the vincentiles of the grammatical conditions in these two experiments showed opposite patterns. Thus, our results support the hypothesis that the processing difficulty seen for plural attractors in Experiment 3A may have been due to a Type 1 error. Further studies and possibly the use of different auxiliary constructions will be needed to settle the question of whether attraction effects exist in grammatical sentences in Spanish. One problem with the current experiments is that the future auxiliary constructions always contained the preposition "a" as the immediate post-verbal region. As this word is short, very frequent and it invariably appeared in the same position, it may have become quite predictable for some participants, resulting in many trials with extremely fast reaction times. This may have increased variability across participants' responses, which may have contributed to the contradictory profiles that we found for grammatical sentences.

General discussion

The current experiments investigated the question of how morphological variation impacts language comprehension by comparing the processing of subject–verb

agreement in Spanish and English. We considered a key operation in comprehension, the retrieval of previous information from memory, and proposed several ways in which this mechanism could be affected by a richer morphological system. Our results support the claim that Spanish and English speakers implement retrieval uniformly, and they suggest that attraction errors in comprehension arise due to an error-driven mechanism. These claims are discussed below.

Attraction across languages

We adopted an account under which comprehenders experience attraction effects when they erroneously retrieve a number-matching attractor noun from memory. The comparison between Spanish and English provides insight into how a richer and functionally more relevant morphology can affect the retrieval mechanism. We expected smaller or null attraction effects if Spanish comprehenders were more certain about their encoding of number information prior to the verb and they engaged less in retrieval as a reanalysis mechanism. Alternatively, larger attraction effects were expected if the morphological cues of the verb were weighted more strongly at retrieval and resulted in more incorrect retrievals of the attractor noun in Spanish. In contrast, if the retrieval mechanism was engaged uniformly across languages, irrespective of surface morphological differences, quantitatively similar attraction effects were expected across Spanish and English.

Our results rule out a strong version of the first hypothesis, as Spanish comprehenders showed robust attraction effects in every experiment. However, a full comparison of these three hypotheses should also address whether the size of attraction effects was similar across languages. Cross-linguistic comparisons are challenging because they necessarily involve comparing across experiments with different participants and items. The data plotted in Fig. 9 provides a starting point, by comparing the effects of grammaticality and attraction in the Spanish and English experiments, which were conducted with similar participant populations and similar sentence materials. This was particularly true in the case of auxiliaries, as the items from Experiment 3B were a direct translation of the items from English Experiment 2. In the plot, the effect of grammaticality reflects the amount of slowdown due to ungrammaticality, while the effect of attraction shows the amount of facilitation in the ungrammatical conditions due to the presence of a plural attractor.[3] Mean effect sizes were computed using a standardized scale, Cohen's *d*, where the difference between the means in the two conditions of interest is divided by the pooled standard deviation (Cohen, 1969; Kirby & Gerlanc, 2013). 95% CIs were computed by bootstrapping.

[3] The effect of grammaticality was calculated as the mean RT difference between ungrammatical and grammatical sentences in the no attractor conditions ($Grammaticality = \overline{X}_{ungram.sgattr} - \overline{X}_{gram.sgattr}$). The effect of attraction was calculated as the mean RT difference between the plural and singular noun conditions in ungrammatical sentences ($Attraction_{ES} = \overline{X}_{ungram.sgattr} - \overline{X}_{ungram.plattr}$).

Two useful generalizations emerge from Fig. 9. First, a within-language comparison of grammaticality and attraction effects reveals that while in English these effects are closer to parity, in Spanish grammaticality effects are always larger than attraction effects. Second, a cross-language comparison reveals that whereas the size of grammaticality effects tends to be larger in Spanish than in English, attraction effects across languages are quite similar. This pattern suggests that the effect of a richer morphology is to make comprehenders more sensitive to agreement violations overall, but that the amount of facilitation due to attraction is quantitatively similar across languages.

Why should attraction effects be similar across languages? Previous well-known accounts have proposed that grammatical cues are weighted differentially in processing depending on their availability in a language (e.g. the Competition Model: MacWhinney, 1987; MacWhinney & Bates, 1989). For example, in tasks that require agent identification in the presence of conflicting cues, English speakers overwhelmingly use word order and assign agenthood to sentence-initial nouns, while German and Italian speakers show more sensitivity to other cues such as agreement and animacy (Bates, McNew, MacWhinney, Devescovi, & Smith, 1982). Given these previous findings, it is striking that attraction effects were not influenced by the increased availability of number morphology in Spanish.

We suggest that the similarity in the size of attraction effects in Spanish and English might be due to the role of retrieval as a repair mechanism in comprehension more generally, although the current study focused on its use in agreement computation. As speakers of all languages have to deal with noisy, sometimes erroneous input, repair strategies might be universal, and thus less affected by language-specific grammatical properties. Further research is necessary to address this hypothesis, but we suggest that a useful dimension to predict where languages should (and should not) differ in comprehension is to consider the function of the cognitive mechanisms under study, and whether they are used during first pass vs. error-driven processing.

Lastly, our results are consistent with the cross-linguistic production literature on agreement attraction. Previous results have been mixed, as some have found higher attraction rates in languages with poor morphology, whereas others have found the opposite pattern (see Lorimor et al., 2008, for discussion). However, a recent study conducted a more controlled comparison across languages by using lexically matched materials in Spanish and English and manipulating whether sentence preambles allowed for a distributive reading (e.g. "The label on the bottles" vs. "The key to the cabinets"; Bock et al., 2012). They found that English and Spanish speakers produced a similar number of plural verbs when the preambles had a plural attractor (*Spanish*: 37 responses, 7.4%; *English*: 38 responses, 7.2%) and that distributivity increased the likelihood of plural responses similarly across languages (*Spanish*: 6%; *English*: 9%).

A question for further research is whether grammatical sentences ever give rise to attraction effects in comprehension. While these effects are generally absent in English

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002

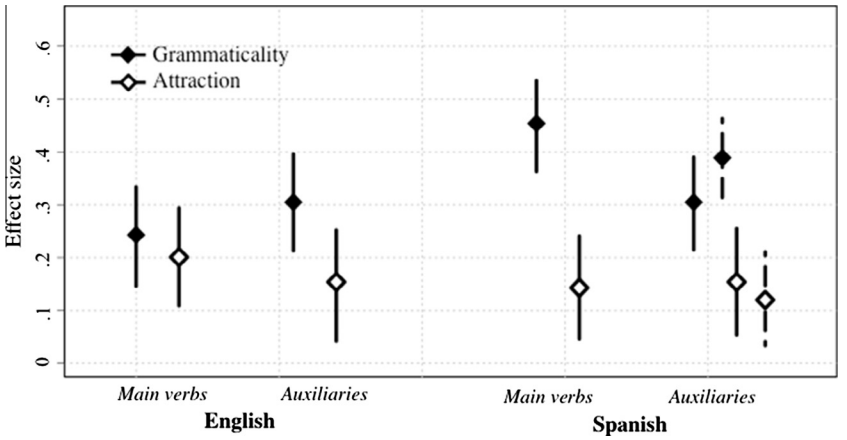

Fig. 9. Comparison of grammaticality and attraction effect sizes across four experiments using Cohen's *d* scores. The effect of grammaticality shows the mean RT difference between ungrammatical and grammatical sentences in the no attractor conditions. The effect of attraction shows the mean RT difference between the singular and plural noun conditions in ungrammatical sentences. Diamonds show mean effect sizes, and vertical bars display bootstrapped 95% CIs. Effect sizes were always estimated from length-regressed RTs in the post-verbal region. Data for English main verbs was obtained from WL&P (2009: Experiment 2). For Spanish auxiliaries, solid lines correspond to Experiment 3A and dashed lines to Experiment 3B.

and were also absent with Spanish main verbs (Experiment 1), we observed attraction with Spanish auxiliaries in grammatical sentences in Experiment 3A. Grammatical attraction effects have also been observed in a recent eye-tracking experiment in Spanish, where sentences with plural attractors elicited longer reading times and less regressions to verbs (Acuña-Fariña et al., 2014). However, the attraction effect in our first auxiliary experiment was only observed in one region and it did not replicate in Experiment 3B. Thus, we cannot rule out that the possibility that the effect in Experiment 3A was due to a Type I error. Overall, what we can conclude is that whereas attraction effects in ungrammatical sentences are robust and obtain reliably across studies and languages, attraction effects in grammatical sentences are less stable. More comprehension studies are needed to address the reasons for this contrast.

A two-stage mechanism for attraction effects in comprehension

The vincentile plots provide two useful insights about attraction effects in comprehension. The first is that attraction effects only take place in a minority of trials, namely, those that correspond to the longest reaction times, which are indexed by the late vincentiles. The second is that grammaticality effects appear on earlier vincentiles than attraction effects, consistent with an account in which attraction follows the initial disruption caused by the detection of an agreement violation.

The finding that attraction only influences a small number of trials replicates previous findings by Staub (2009, 2010), who used similar RC constructions in a forced-choice paradigm in which participants read sentence preambles and made speeded choices between singular and plural verbs. By modeling the latency of correct responses using the ex-Gaussian distribution, Staub found that increases in latency due to attraction were carried by only a small subset of the reaction times, with little or no slowing on other trials. Here we found a similar pattern using a more naturalistic reading paradigm. However, the comparison across studies should be made cautiously due to several experimental differences. The first is that the nature of our task did not allow us to split trials into correct and incorrect responses, as participants were not asked to make conscious decisions. Therefore, our reading profiles reflect a mixture of both types of responses and should not be mapped directly to previous studies that separated those responses. Second, since previous sentence processing experiments have rarely had enough items per condition to be examined with vincentile analyses, we cannot rule out the possibility that properties of the self-paced reading task itself would drive any reading time effect to show up selectively on later vincentiles.

The second finding from the vincentiles is that grammaticality effects generally showed an earlier onset than attraction effects. This pattern is consistent with the idea that attraction mainly takes place in trials where comprehenders are already experiencing processing difficulty due to the encounter of an agreement violation. This contrasts with accounts that assume that retrieval takes place in both grammatical and ungrammatical sentences and also with percolation accounts, which claim that attraction is due to the incorrect encoding of the number of the subject phrase, such that comprehenders sometimes fail to notice the ungrammaticality of the sentence and effects of grammaticality and attraction are simultaneous.

At the same time, it is important to point out that while vincentile analyses provide useful additional information to the mean reaction time analysis, their interpretation is subject to several limitations. The first is that vincentiles cannot provide an absolute timing estimate for effects of grammaticality and attraction. This is because the vincentiles were computed in the verb+1 region, and thus they do not uniformly map onto the time elapsed since the agreement information was first encountered at the verb: readers may have spent varying amounts of time reading the verb region before moving on to the spillover region.

Thus, although the vincentiles cannot provide absolute timing estimates, they suggest that by the time when attractor number impacts the vincentiles, readers had already shown evidence of processing difficulty associated with the ungrammaticality of the sentence.

A second limitation is that the onset differences between grammaticality and attraction effects in the vincentiles cannot be taken as definite evidence that the underlying cognitive processes began at different times. Whereas the earlier onset of grammaticality is consistent with the claim that the cognitive processes associated with grammaticality unfolded earlier than the processes associated with attraction, the same vincentile effect is also consistent with the two types of processes having a matched time-course but overall differences in strength. In other words, attraction could have affected the vincentiles as early as grammaticality but less strongly, leading the statistical analysis to find a significant difference only at a later vincentile. In the current studies we favor an interpretation based on time course differences, because this is also consistent with the pattern of the reading time means in Experiments 1 and 3B, where grammaticality effects were observed immediately at the verb region but attraction was only observed at the following region. However, further work will be needed to more effectively distinguish between an effect-strength and an onset latency interpretation of these vincentile time-course differences.

In sum, the vincentile plots are consistent with the hypothesis that participants first detect an agreement violation, and then initiate a cue-based retrieval process in search of a plural noun to license the verb. In some cases, this retrieval results in the selection of a syntactically inappropriate but number-matching noun, giving rise to attraction. We suggest that an important question for future research is whether this two-stage model applies to other dependencies (e.g., negative polarity licensing) that have been previously linked to a retrieval mechanism (Vasishth, Bruessow, Lewis, & Drenhaus, 2008).

Conclusion

We have shown that despite clear differences in the richness and functional importance of number morphology in Spanish and English, Spanish comprehenders are susceptible to the same attraction errors in subject–verb agreement as English comprehenders, and that the magnitude of these effects is strikingly similar across languages. Furthermore, the distributional profile of attraction effects in Spanish is consistent with a two-stage process in which attraction errors are due to a retrieval mechanism that is triggered by the initial detection of an agreement violation. Together these data suggest that memory retrieval is implemented similarly across languages in comprehension.

Acknowledgments

We thank Jonathan Simon, Brian Dillon, Ewan Dunbar, Yuan Liao, Matt Wagers, Bill Idsardi and Diogo Almeida for statistical advice, and Eleonora Pinotti, Carolina Gattei, María del Mar Contaldi and Tim Dawson for their invaluable help running the experiments. This work was supported in part by NSF grant BCS-0848554 to Colin Phillips.

A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jml.2015.02.002.

References

Acuña-Fariña, J. C, Meseguer, E., & Carreiras, M. (2014). Gender and number agreement in comprehension in Spanish. *Lingua, 143*, 108–128.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283–316.

Balota, D. A., Yap, M. J., Cortese, M. I., & Watson, J. M. (2008). Beyond response latency: An RT distributional analysis of semantic priming. *Journal of Memory and Language, 59*, 495–523.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7*. <http://CRAN.R-project.org/package=lme4>.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition, 11*, 245–299.

Bock, J. K., Carreiras, M., & Meseguer, E. (2012). Number meaning and number grammar in English and Spanish. *Journal of Memory and Language, 66*, 17–37.

Caplan, D., & Waters, G. (2013). Memory mechanisms supporting syntactic comprehension. *Psychological Bulletin and Review, 20*, 243–268.

Clifton, C., Jr., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica, 11*, 11–39.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic.

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting interference profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language, 69*, 85–103.

Eberhard, K. M., Cutting, J. C., & Bock, J. C. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review, 112*, 531–559.

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. H. (2008). The interplay of syntax and form in language production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes, 23*, 329–374.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Attraction in sentence production: The role of syntactic structure. *Language and Cognitive Processes, 17*, 371–404.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 1411–1423.

Hartsuiker, R. J., Schriefers, H. J., Bock, J. K., & Kikstra, G. M. (2003). Morphophonological influences on the construction of subject–verb agreement. *Memory & Cognition, 31*, 1316–1326.

Hofmeister, P. (2010). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes, 26*, 376–405.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

Jia, G., Aaronson, D., & Wu, Y. (2002). Long-term language attainment in bilingual immigrants: Predictive variables and language group differences. *Applied Psycholinguistics, 23*, 599–621.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603–634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*, 1–33.

Just, M. A., & Carpenter, P. A. (1978). Inference processes during reading: Reflections from eye fixations. In J. W. Senders, D. F. Fisher, & R. A.

Monty (Eds.), *Eye movements and the higher psychological functions*. Hillsdale, New Jersey: Erlbaum.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 3*, 228–238.

Kail, M. (1989). Cue validity, cue cost, and processing types in French and Spanish. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of language processing* (pp. 77–117). New York: Cambridge University Press.

Kirby, K. N., & Gerlanc, D. (2013). BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior Research Methods, 45*, 905–927.

Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2014). *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0–11*. <http://CRAN.R-project.org/package=lmerTest>.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*, 375–419.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10*, 44–54.

Lorimor, H., Bock, J. K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes, 23*, 769–799.

MacWhinney, B., & Bates, E. (Eds.). (1989). *The cross-linguistic study of sentence processing*. New York: Cambridge University Press.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior, 23*, 127–150.

MacWhinney, B. (1987). Cues, competition, and learning. In J. Miller (Ed.), *Wisconsin papers on language development*. Madison, Wisconsin: University of Wisconsin.

MacWhinney, B. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction*. New York: Cambridge University Press.

Mancini, S., Molinaro, N., & Carreiras, M. (2013). Anchoring agreement. In M. Sanz, M. Tanenhaus, & I. Laka (Eds.), *Language down the garden path*. Oxford: Oxford University Press.

McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21*, 395–423.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language, 48*, 67–91.

Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language, 36*, 569–587.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language, 41*, 427–456.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446–461.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*, 510–532.

Severens, E., Jansma, B., & Hartsuiker, R. (2008). Morphophonological influences on the comprehension of subject–verb agreement: An ERP study. *Brain Research, 1228*, 135–144.

Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language, 60*, 308–327.

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition, 114*, 447–454.

Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review, 18*, 371–376.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1280–1293.

Tanner, D., Nicol, J., & Brehm, L. (2014). The time course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language, 76*, 95–215.

Vasishth, S., Bruessow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science, 32*, 685–712.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject–verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition, 61*, 261–298.

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject–verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language, 34*, 186–215.

Vigliocco, G., Hartsuiker, R. J., Jarema, G., & Kolk, H. H. J. (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes, 11*, 407–442.

Vincent, S. B. (1912). The function of the viborissae in the behavior of the white rat. *Behavioral Monographs, 1*, 5.

Wagers, M., Lau, E., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language, 61*, 206–237.

Please cite this article in press as: Lago, S., et al. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* (2015), http://dx.doi.org/10.1016/j.jml.2015.02.002