Evelyn Fernández-Lizárraga
LING 245B
June 8, 2021

Spanish Definite Articles with Proper Names: The Role of Common Ground[1]

1 INTRODUCTION

The present study investigates the relation between shared common ground and the use of the Spanish definite article with proper names (DA+NAME; e.g. La María, El Juan, etc.) through an online experiment that tests speakers' judgments regarding felicitous uses and affective perceptions of the DA+NAME construction embedded in neutral contexts.

Variation in the use of definite determiners with proper names appears to be a widespread phenomenon in Spanish. Usage and social perceptions vary across regional dialects, with socioeconomic status, frequency of use, and the gender of the referent playing a key role in framing language attitudes (Tieperman, 2020). Where this variation occurs, there are generally two broad interpretations. The definite article may signal positive affect toward the referent (e.g. affection, familiarity, closeness, etc.), or conversely, it may indicate a negative opinion of the referent (e.g. regarded as condescending or insulting, distancing, objectifying, etc.). The present study is an initial attempt into understanding how speakers arrive at such disparate interpretations of the same phenomena. Critically, I focus on the role of common ground and felicity, as well as baseline affective interpretations of the DA+NAME construction.

2 BACKGROUND

Christodulelis (2017) argues that use of the article conveys two types of implicature (Christodulelis, 2016, 2017). First, the article signals a conventional implicature, "indicating the speaker's non-neutrality towards the referent" (cf. Potts, 2007). Second, the article also conveys a particularized conversational implicature, thus "with context indicating the precise flavor of the speaker's attitude." Using an online quantitative experimental study open to any native speaker of Spanish, participants were presented with contexts that framed the speaker's attitude towards the referent as either positive, negative, or neutral. Participants rated each target sentence along two different Likert scales, an acceptability scale and an intensity scale. The use of definite articles was less felicitous in neutral contexts. In cases that involved the definite article (in contrast to the

[1] For experiment materials: https://github.com/lfre94/spanish-definite-articles
For OSF Preregistration: https://osf.io/z5bh7

same name presented in the same context without the definite article), the participants ranked the speaker as having a stronger attitude towards the referent (participants only ranked the "intensity" of the attitude but not whether the attitude was positive or negative).

However, it appears that the neutral contexts in Christodulelis (2017) include what she calls "out of the blue" contexts, situations in which the "utterance [with the referent] was produced with little introduction" (2017, p. 11), that is, the hearer did not know the referent. These "out of the blue" contexts were found to be highly infelicitous in initial qualitative work on Mexican Spanish (Christodulelis, 2016). While Christodulelis argues that "the hearer must know of the referent sufficiently well (i.e there must be sufficient material in his/her common ground with the speaker) to draw the intended inference of the speaker's attitude" (2017, p. 11), from the example stimuli provided and description of the experimental design, it remains unclear whether the neutral contexts accounted for common ground factors.

If the neutral contexts did not clearly specify whether the referent was in common ground, the low acceptability ratings for the neutral scenarios in Christodulelis (2017) may be cases that are also infelicitous for unmarked proper names. As Prince (1992) argues, proper names may be discourse-new or discourse-old but must always be hearer-old in order to be felicitous. In other words, proper names can be introduced into discourse with no introduction so long as they already exist in the interlocutors' common ground.

The present study takes inspiration from Christodulelis (2017), using a similar experimental design to test whether the status of a referent as hearer-old is sufficient for the use of the definite article to be deemed felicitous. This study focuses on the role of common ground in *neutral* contexts, that is, scenarios in which participants are not given explicit information as to whether the speaker and referent have a positive or negative relationship (cf. Christodulelis, 2017). Thus, the key research questions are the following:

1. In what contexts is use of the definite article with a proper name (DA+NAME) felicitous? More specifically, what is the role of shared common ground between interlocutors?
2. What are speakers' judgments regarding affective uses of the DA+NAME construction?

Following Christodulelis (2017), the present study tests the relationship between the marked conditions and felicity as well as affect by asking participants to rate target sentences along two separate scales. For the first research question, the prediction is that participants will rate DA+NAME constructions as more felicitous in contexts where the referent is in common ground between the speaker and listener. Given the results of both Christodulelis (2017) and Tieperman (2020), the prediction is that participants will rate DA+NAME constructions at the extremes of the affect scale (as either very positive or very negative) in comparison to the unmarked form without the definite article. However, this is

also the first study to provide participants with only neutral contexts, rather than overt positive or negative scenarios. While no context is truly neutral, by including an affect scale, this study gauges baseline interpretations of the definite article.

3 EXPERIMENT

3.1 Methods

Participants I recruited 200 participants on Amazon's Mechanical Turk. The experiment was open to any adult native speaker of Spanish. Participants were paid $5 for completing the experiment, using a rate of $15 per hour. Participants who achieved less than 80% accuracy on attention checks (N = 26) or were non-native Spanish speakers (N = 39) based on self-reported demographic data, were excluded from the analysis. Given the length of the experiment, an additional non-preregistered exclusion included participants who completed the experiment in less than 10 minutes (N = 9; the mean completion time was 21.47 minutes). Thus, the final analysis included 131 participants.

Since demographic data regarding Spanish turkers remains quite limited (Ortega-Santos, 2019), I include here additional participant social (Table 1) and by country information (Table 2) that may be of interest to future work in Spanish sociolinguistics.

Self-Reported Demographic Information		
Gender	Female	56
	Male	74
	Non-Binary	1
Age	18-63; Mean 33	
Education	Elementary	1
	Middle School	7
	High School	37
	College	62
	Graduate	22
	Prefer not to answer	2

Table 1: Participant demographic information

Participants by Country	
Argentina	17
Bolivia	1
Brazil	7
Chile	2
Colombia	7
Ecuador	3
El Salvador	1
Mexico	16
Nicaragua	1
Puerto Rico	1
Spain	21
Uruguay	3
USA	35
Venezuela	14
Prefer not to answer	2

Table 2: Participant country information

Procedure Before beginning the experiment, participants were asked to verify that they were a native speaker of Spanish by listening to an audio recording in Spanish and typing the correct response.[2] Participants then completed three example trials before continuing to the main experiment. The first trial demonstrated an example of a syntactically correct but pragmatically odd sentence; the second trial demonstrated a referential term with high positive affect; and the third trial demonstrated a referential term with high negative affect and introduced attention checks. Once completed, participants proceeded to the critical trials, which had the same format as the example trials (i.e., a dummy context followed by a target sentence in blue and slider scales). For each critical trial, participants were given a dummy context that involved three entities — a speaker, a listener, and a referent. Each context began with a sentence about the speaker and concluded with the speaker addressing the listener, with the target sentence in quotes and marked in blue. The context specified the relationships between all three entities, explicitly stating whether the listener knows the referent. In order to help participants keep track of these fictional relationships, only the speaker and referent were referred to by names; listeners were referred to by their relation to the speaker (e.g. "sister," "friend," "colleague," etc.). Figure 1 provides an example of a critical trial —

[2] Compared to the pilot data, using audio stimuli was more effective than text questions in limiting non-native speakers (and non-human participants) from taking the experiment. However, a high number of participants still had to be excluded. Thus, future work should consider randomizing questions in addition to using audio or visual stimuli.

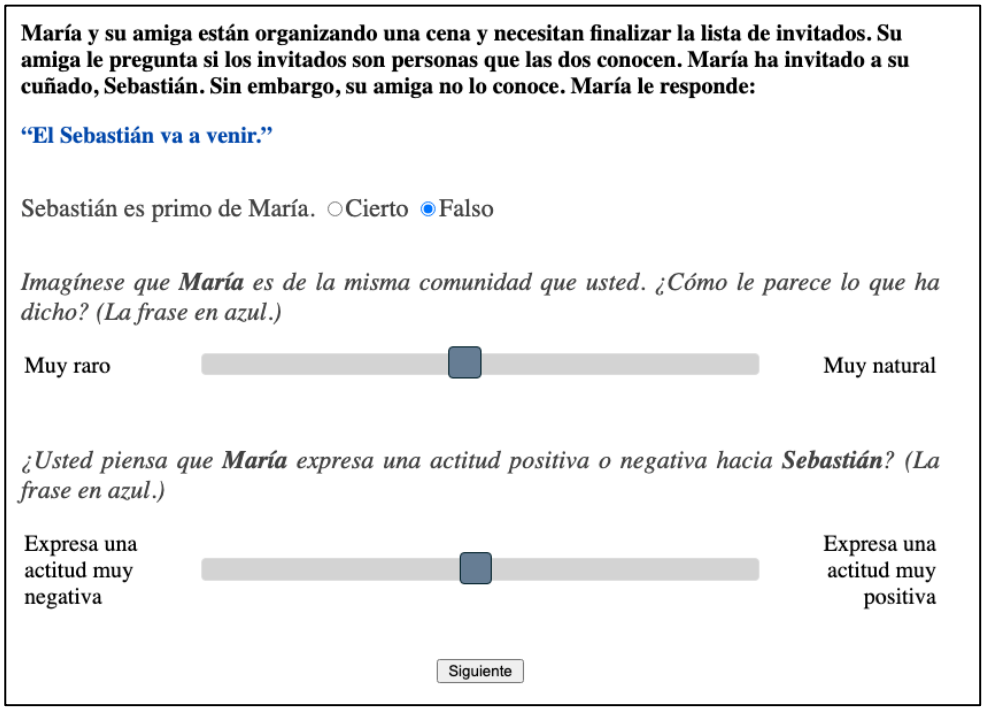

Figure 1: Example of critical trial — Item Type: Dinner; Context Condition: No Common Ground; DA presence: DA. (Context translation provided in Table 3)

After reading the context and the target sentence, participants answered a simple true or false question about the context as an attention check (e.g. in Figure 1, "Sebastián is María's cousin." The correct answer is false since he is her brother-in-law). Participants were then asked to rate the target sentence in blue along two slider scales. The first question asked, "Imagine that **María** [speaker name] is from the same community as you. What do you think of what they have said? (The phrase in blue.)" Participants rated the sentence as either "very strange" or "very natural." The second question asked, "Do you think **María** [speaker name] expresses a positive or negative attitude towards **Sebastián** [referent name]? (The phrase in blue.)" Participants rated the sentence as either "[It] expresses a very negative attitude" or "[It] expresses a very positive attitude." Participants had to complete the attention check and the two slider scales before continuing to the next trial.

The format of the trials differed from Christodulelis (2017) in two critical ways. First, the present study used slider scales rather than a 5-point Likert scale. Secondly, Christodulelis asked participants to rate the target sentence along an intensity scale rather than affect scale (e.g. "How strongly does the speaker convey his/her feelings toward the person that he/she is talking about?"). Since I only presented participants with neutral contexts, lacking overt positive or negative scenarios, the affect scale served as method of capturing participants baseline affective inferences regarding the DA+NAME construction.

After completing the trials, participants completed a short demographic survey. In addition to the information provided above, participants also self-reported how often they use the variant and how often they hear the variant being used in their community. Participants had the option of describing when they use definite articles with names as well commenting on their perceptions of the variant.

Materials While I retained the basic format of Christodulelis (2017) (i.e. a dummy context and target sentence, followed by an acceptability and affect scale), the rest of the experimental design was significantly different.

Participants were presented with 8 critical trials and 8 filler trials. The critical trials consisted of 8 item types. For each item type, there were two possible context conditions. Participants randomly received either 1) a neutral context with no common ground (i.e. the listener does not know the referent); or 2) a neutral context with common ground (i.e. the listener knows the referent). For each item type, there were also two possible target sentence conditions. Each context was randomly paired with a neutral carrier sentence that either contained the DA or did not. This information is summarized in the table below with examples —

Context Condition	**Sample Context (Item Type: Dinner)**	**DA Presence (Target Sentence)**
No Common Ground	María y su amiga están organizando una cena y necesitan finalizar la lista de invitados. Su amiga le pregunta si los invitados son personas que las dos conocen. María ha invitado a su cuñado, Sebastián. Sin embargo, su amiga no lo conoce. María le responde:	"El Sebastián va a venir."
		"Sebastián va a venir."
	María and her friend are organizing a dinner and (they) need to finalize the list of guests. Her friend asks her if the guests are people they both know. María has invited her brother-in-law, Sebastián. ***Nevertheless, her friend does not know him.*** *María responds (to her friend):*	*"**The** Sebastián is coming."*
		"Sebastián is coming."
Common Ground	María y su amiga están organizando una cena y necesitan finalizar la lista de invitados. Su amiga le pregunta si los invitados son personas que las dos conocen. María ha invitado a su cuñado, Sebastián, a quien su amiga conoce. María le responde:	"El Sebastián va a venir."
		"Sebastián va a venir."

	María and her friend are organizing a dinner and (they) need to finalize the list of guests. Her friend asks her if the guests are people they both know. María has invited her brother-in-law, Sebastián, ***whom her friend knows.*** *María responds (to her friend):*	*"**The** Sebastián is coming."*
		"Sebastián is coming."

Table 3: Experimental Design

Participants all viewed the same 8 fillers. These fillers followed the same format as the critical context conditions; they consisted of a dummy context with information about the common ground followed by a target sentence. Six of the fillers contained neutral contexts. Of the two remaining fillers, one consisted of a very positive scenario and the other of a very negative scenario; these two fillers serve to anchor the affect scale in the critical trials. Regarding the target sentences, referent names were marked using either diminutives, demonstratives (e.g. *Esa/Ese* 'that'), or common terms of address (e.g. *Doña/Don, Sra./Sr.* 'Mrs./Mr.'*).* Across critical and filler trials, names for speaker and referent were controlled for gender. That is, there was an even distribution of masculine speakers paired with masculine referents; feminine speakers paired with feminine referents; feminine speakers paired with masculine referents; and masculine speakers paired with feminine referents.

3.2 Results

After exclusions, the data consisted of 1,048 observations across all critical trials. Due to experimenter error, conditions paired with DA target sentences were oversampled. Thus, there are more data points for DA target sentence conditions (with Common Ground, N = 346; with No Common Ground, N = 319) compared to the No DA sentence conditions (with Common Ground, N = 194; with No Common Ground, N = 189). In this section, I begin by summarizing the results for felicity ratings followed by affect ratings.

Figure 2 demonstrates mean felicity ratings by DA presence for Common Ground and No Common Ground context conditions. To test whether DA presence and context conditions affected slider ratings for felicity, I conducted a mixed-effects linear regression predicting felicity from a centered presence of the DA presence predictor (reference level = DA) and a centered context condition predictor (reference level = No Common Ground), and their interaction.[3] I included by-subject and by-item random intercepts as well as random slopes for DA presence and context conditions. There was a main effect of DA presence, such that target sentences with No DA presence received significantly higher ratings for felicity (β = 0.11, SE = 0.02, t = 6.46, $p < 0.0001$). Additionally, there was a main

[3] All analyses were conducted in R (R Core Team, 2020) using the *lme4 package* (Bates et al., 2015).

effect of context condition, such that target sentences which occurred in Common Ground contexts also received higher felicity ratings (β = 0.07, SE = 0.02, t = 4.73, p < 0.001). Furthermore, the interaction between DA presence and context condition was also significant. With respect to context conditions, simple effects analysis revealed that the interaction is driven by the No DA conditions undergoing a greater change in felicity ratings by context condition (i.e. No DA with Common Ground rated higher than No DA with No Common Ground; β = 0.20, SE = 0.03, t = 7.13, p < 0.0001) as compared to smaller changes in DA conditions by context condition (i.e. DA with Common Ground rated higher than DA with No Common Ground ; β = 0.11, SE = 0.02, t = 4.935, p < 0.0001). A second simple effects model analyzed the interaction with respect to changes in DA presence within context conditions. This analysis demonstrated that there was a greater difference in felicity ratings for DA presence within the Common Ground conditions than within the No Common Ground conditions. That is, Common Ground conditions with No DA were rated higher in felicity than DA conditions with Common Ground (β = 0.26, SE = 0.04, t = 7.11, p < 0.0001). This change in ratings was greater than a similar change observed in the No Common Ground conditions with No DA as compared to the No Common Ground with DA conditions (β = 0.17, SE = 0.04, t = 4.49, p < 0.001).

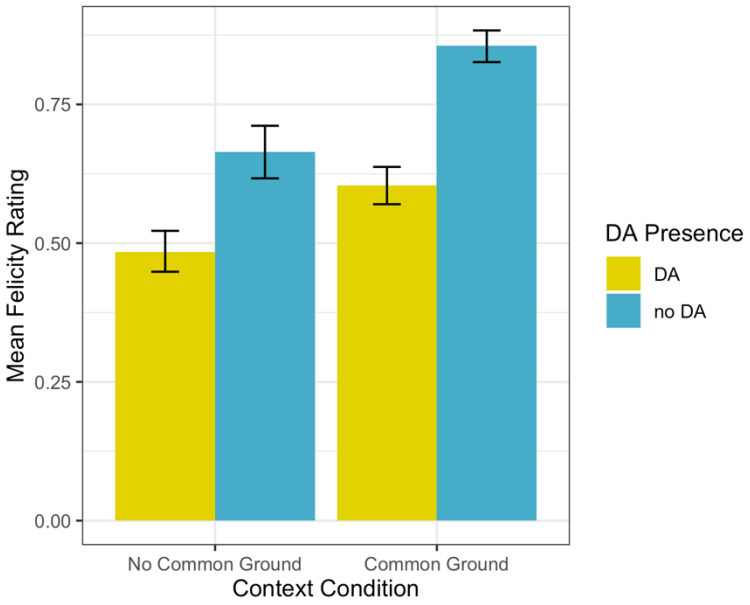

Figure 2: Mean felicity ratings by DA presence for Common Ground and No Common Ground context conditions.

Turning to affect, Figure 3 demonstrates mean affect ratings by DA presence for Common Ground and No Common Ground context conditions. To test the relationship between affect slider ratings, DA presence and context condition, I conducted a mixed-effects linear regression predicting affect from a centered presence of the DA presence

predictor (reference level = DA) and a centered context condition predictor (reference level = No Common Ground), and their interaction. I included by-subject random intercepts as well as random slopes for DA presence and context conditions. For item type, I only included by-item random intercepts as this was the maximal random effects structure justified by the design. The analysis revealed a main effect of DA presence (β = 0.03, SE = 0.01, t = 3.85, p < 0.001), such that target sentences with No DA were rated higher than DA conditions. Furthermore, Context Condition was also a significant predictor (β = 0.02, SE = 0.01, t = 3.34, p < 0.01), such that target sentences in the Common Ground conditions were rated higher than No Common Ground conditions. The interaction between DA presence and context condition was not significant.

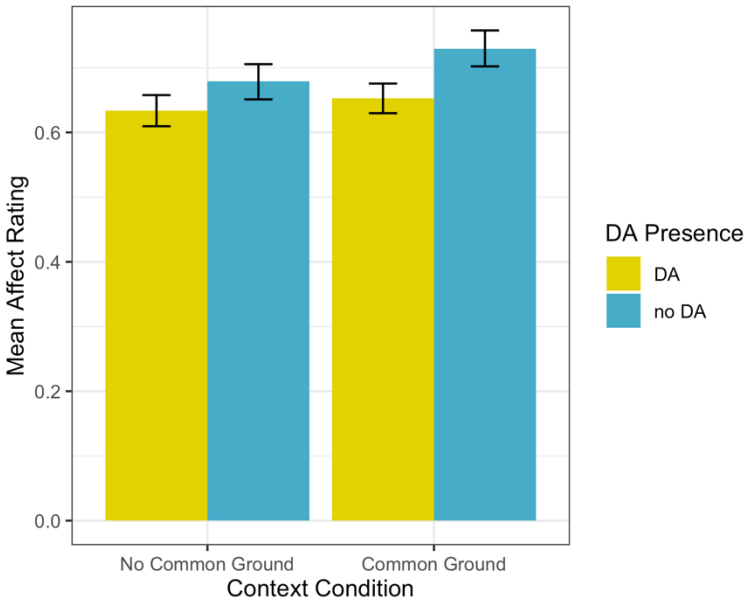

Figure 3: Mean affect ratings by DA presence for Common Ground and No Common Ground context conditions.

4 Discussion

This study pursued two main research questions, predicting felicity and affect slider ratings as a function of the DA presence and context condition. Regarding felicity, the main prediction was borne out; participants rated DA+NAME as more felicitous when the referent was in common ground between the speaker and listener. Even so, as shown in Figure 2, common ground plays an important role for all uses of proper names since context condition also increased mean felicity ratings for the No DA conditions. Considering the results from Christodulelis (2017), the mean felicity ratings for the DA conditions appear to be higher in the present study. These results are not directly comparable since Christodulelis (2017) had a different experimental design, used Likert scales instead of

slider scales, and reported normalized mean acceptability ratings. Even so, it is worth noting that in the 2017 study, on a scale of -1 (as very strange) and 1 (as very natural), No DA neutral contexts received an average rating of 0.51, and DA neutral contexts received an average rating of -0.64. Since Christodulelis argues that DA+NAME is not felicitous in neutral contexts, this predicts that the present study would have very low felicity ratings for all DA conditions since only neutral contexts were used, which lacked overt affective information. Instead, when common ground conditions are accounted for, some uses of the DA approximate felicity ratings for No DA conditions, even in neutral contexts. For instance, the mean felicity rating for the DA+Common Ground condition was 0.60 which is comparable to the average rating for the NoDA+No Common Ground condition at 0.66 (see figure 3). Taken together, these results suggest that a comprehensive account of DA+NAME variation must consider the properties of proper names (Prince, 1992), particularly the role of common ground. While common ground status may be sufficient to make the definite article felicitous in certain contexts, the present study does not answer why there is a significant difference in felicity ratings for DA presence *within* context conditions. Potentially, it is here where we may see a difference in felicity ratings if overtly positive and negative contexts were added to the experiment.

Turning to the second research question, participants were asked to rate target sentences along an affect scale in order to analyze affective uses of the DA+NAME construction. The main prediction was that participants would rate DA+NAME constructions at the extremes of the affect scale (as either very positive or very negative) in comparison to the unmarked form without the definite article. This prediction was borne out in part. First, the mixed effects analysis demonstrated that DA presence was a significant predictor for affect ratings, such that No DA conditions received a higher affect rating. Even so, the mean affect ratings across conditions were not strikingly different (see figure 3). However, these averages obscure some of the variability across item types, as shown in figure 4, which demonstrates felicity and affect ratings for DA presence across critical item types. The original hypotheses predicted a U-shape curve for DA presence, such that target sentences rated high or low on the affect scale would also receive a high felicity rating, while target sentences rated at 0.5 on the affect scale would receive a low felicity rating. The item type "dinner" potentially demonstrates this pattern, since the extremes of the scale have an upwards slope. Figure 4 also shows that it is generally the DA conditions which encompass the full range of the affect scale, while the No DA conditions tend to be rated 0.5 or higher on the affect scale.

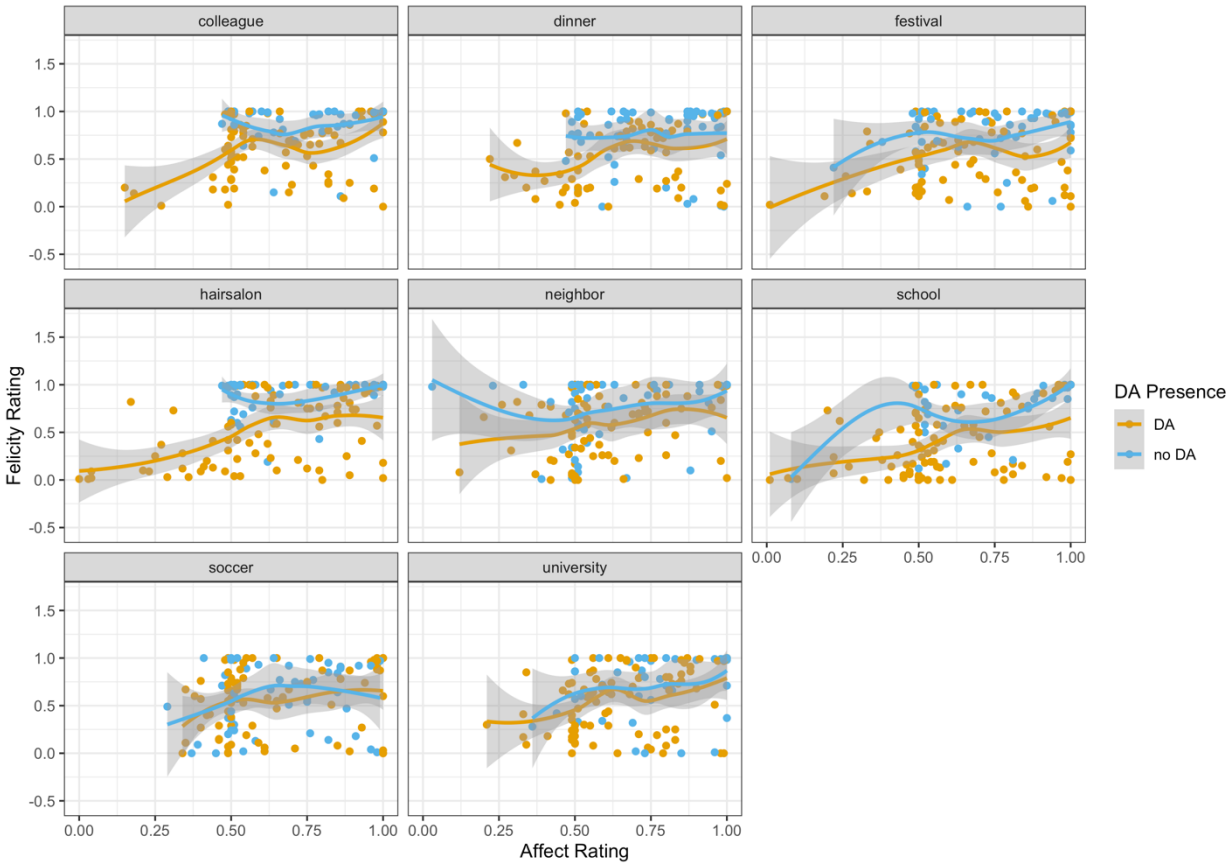

Figure 4: Felicity and affect ratings for DA presence by item type for critical trials

There are several potential reasons for variability across item type which merit further investigation[4]. The professional or businesslike settings ("colleague," "hairsalon," "school") display a wide range in affect. Participants who use the variant reported in the post-experiment survey that they only use DA+NAME in informal settings, with only close friends and family, suggesting speech style or register expectations. Additionally, use of DA+NAME may be gendered; it is notable that the lowest affect ratings occurred with female referents (item types "festival," "hairsalon," "neighbor," and "school") or when the speaker was a woman (item types "colleague," "dinner," "hairsalon," "neighbor"). In contrast, the item type "soccer" included a male speaker and referent and the ratings across DA conditions overlap significantly. The above are not definitive conclusions, but suggestions for future research. It is also worth considering whether participants treat the felicity and affect scales as distinct from one another. It appears that positive affect correlates with

[4] While not discussed in this paper, the R analysis file for this experiment (spanish_definite_articles.Rmd) contains additional exploratory analyses and visualizations of the social information data collected in this experiment, such as felicity and affect ratings by age, education, region, frequency of use, and familiarity with the form. Country of origin stands out as a key factor for future studies.

high felicity, while negative affect correlates with low felicity. This trend appears as well in the positive and negative affect fillers in the figure below.

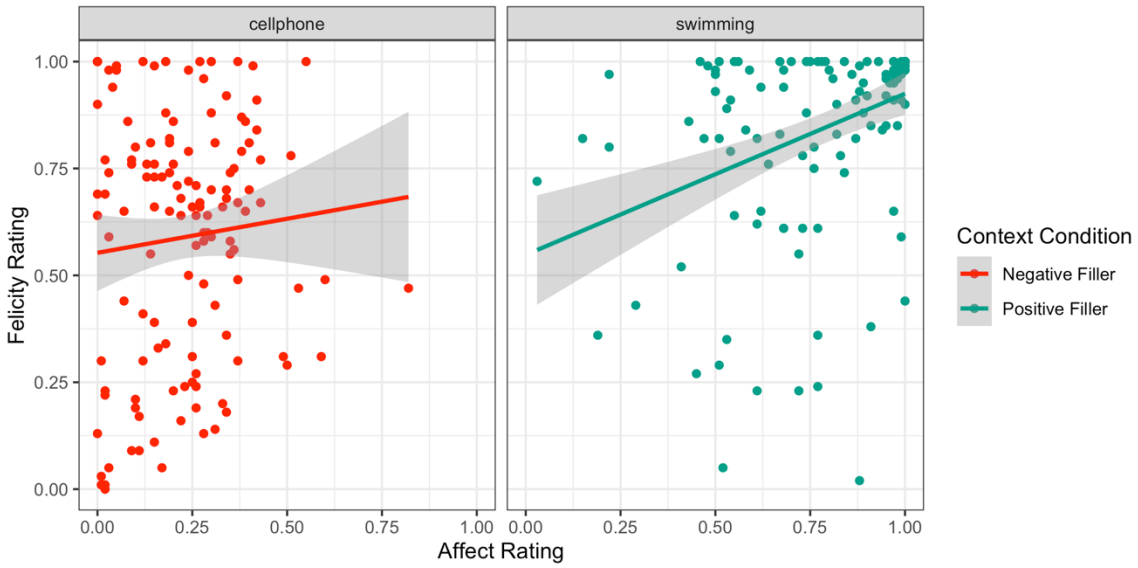

Figure 5: Felicity and affect ratings for negative and positive affect filler trials

5 Conclusion

This experiment is an initial step into the study of definite articles with proper names in Spanish. Importantly, it highlights the role of common ground, demonstrating that the DA+NAME construction may be felicitous in contexts without overt positive or negative affect. This suggests that these definite articles may have other discursive functions aside from affect, such as marking a referent as particularly salient (thus making the DA conditions even less felicitous than the No DA conditions in No Common Ground contexts). While affective uses were not the focus of this experiment, the variability across item types suggests that participants still draw affective inferences despite the supposed context neutrality. What conditions these affective meanings is an area for future research.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Christodulelis, E. (2016). *Definite Articles with Anthroponyms in Mexican Spanish*. OSUCHILL, Ohio State University.

Christodulelis, E. (2017, October 27). *Definite Articles with Anthroponyms Across Varieties of Spanish: Quantitative Support of Conventional and Conversational Implicature*. Hispanic Linguistics Symposium, Texas Tech University, Lubbock, TX.

Ortega-Santos, I. (2019). Crowdsourcing for Hispanic Linguistics and Beyond: Amazon's Mechanical Turk (MTurk) as a source of Spanish data. In *Borealis: An International Journal of Hispanic Linguistics* (Vol. 8). https://doi.org/10.7557/1.8.1.4670

Potts, C. (2007). Into the conventional-implicature dimension. *Philosophy Compass*, *2*(4), 665–679.

Prince, E. F. (1992). The ZPG letter: Subjects, definiteness, and information-status. *Discourse Description: Diverse Analyses of a Fund Raising Text*, 295–325.

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Tieperman, R. E. (2020). *¿ Cómo está el Pablo?: Examining Definite Article Use with Personal Proper Names in Three Regional Varieties of Spanish*.