Effects of animacy and linguistic construction on the interpretation of spatial descriptions in English and Spanish*

JAVIER OLLOQUI-REDONDO

Universidad Complutense de Madrid, Spain

THORA TENBRINK

Bangor University, Wales, UK

AND

ANOUSCHKA FOLTZ

Karl-Franzens-Universität Graz, Austria

ABSTRACT

The languages of the world differ in their use of intrinsic, relative, and absolute reference frames to describe spatial relationships, but factors guiding reference frame choices are not yet well understood. This paper addresses the role of animacy and linguistic construction in reference frame choices in English and Spanish. During each trial of two experiments, adult participants saw a spatial scene along with a sentence describing the location of an object (locatum) relative to another object (relatum) that was animate or human(-like) to varying degrees. The scene presented two possible referents for the locatum, and participants decided which referent the description referred to, revealing which reference frame they used to interpret the sentence. Results showed that reference frame choices differed systematically between languages. In English, the non-possessive construction (*X is to the left of Y*) was consistently associated with the relative reference frame, and the possessive construction (*X is on Y's left*) was associated with the intrinsic

reference frame. In Spanish, the intrinsic interpretation was dominant throughout, except for the non-possessive construction with relata that were not anthropomorphic, animate, or human. We discuss the results with respect to the languages' syntactic repertory, and the notion of inalienable possession.

KEYWORDS: spatial cognition, animacy, reference frames, perspective choice.

1. Introduction

Functioning in space is important for the survival of every species. For humans, this importance is reflected in a rich and complex repertory of spatial language, the use of which is of special interest to linguistics in general and Cognitive Linguistics in particular (Zlatev, 2007). According to Carlson and Covell (2005), the most typical goal of spatial language is to inform somebody of the location of a certain object, and the most effective way to achieve this goal is to describe that object's position in relation to another object whose location is known. Following Tenbrink (2011), this paper uses the terms LOCATUM for the object that needs to be LOCATED, and RELATUM for the object that the locatum is RELATED to in order to describe its position. So in *The cat is in front of the house*, the *cat* (locatum) is being located in relation to the *house* (relatum).

To locate objects, speakers draw on three different types of spatial frames of reference (Levinson 1996, 2003), which allow us to describe spatial relationships between a locatum and a relatum based on a perspective (intrinsic or relative frames of reference) or based on a stable directional system (absolute frame of reference). In an intrinsic frame of reference, the perspective is provided by the relatum's intrinsic features, as in *The cat is in front of the car*, where *front* refers to the front part of the car, or *The cat is in front of me/ you*, where the speaker or hearer serves as relatum and also gives the perspective. In a relative frame of reference, the speaker's and/or listener's perspective is used rather than the relatum's intrinsic features, as in *The cat is in front of the table from my point of view*; here the *table* as relatum does not have (nor need) an intrinsic orientation or perspective. Absolute frames of reference rely on some kind of directional system provided by the interactants' culture or environment (e.g., compass directions), as in *Brighton is south of London*.

Over the past decades, cross-cultural research has identified various factors affecting choice of reference frames. In some cultures, people are constantly aware of the actual (absolute) directions in space, as if they had an inbuilt compass; and some cultures do not seem to use a relative reference system at all (Danziger, 1996; Gaby, 2012; Levinson, 2003). However, the preferences

and choices of reference frames in cultures that use all of the three kinds are still poorly understood. The need for a closer look at factors pertaining to the situation and context in which a spatial reference frame is used, rather than overarching cultural ones, has been repeatedly emphasised, as different studies tend to reveal different preferences within a culture (Tenbrink, 2007). Such factors do not have to be situation-specific; languages often exhibit grammatical and/or usage patterns based on more generic features, such as animacy, dynamics, schematicity, and the like (Talmy, 2000).

In this paper, we compare the relative impact of object properties, such as animacy, and choice of syntactic construction on spatial reference frame choices for the lateral axis (i.e., *left* or *right*) in English and Spanish. These languages differ with respect to the syntactic constructions available for spatial reference. In addition, both languages have structures that are affected by animacy (see Section 1.3). Here we ask to what extent animacy and related features of the relatum influence perspective choice (and, thus, reference frame selection) in differently worded spatial descriptions in these two languages. Consider statements (1) and (2):

(1) The ball is to the right of the chair.
(2) The ball is to David's right.

There are two possible interpretations for each statement, as shown in Figure 1. These interpretations depend on whether the speaker keeps his or her own perspective (i.e., uses the relative frame of reference; see Figure 1, left) or adopts the relatum's perspective (i.e., uses the intrinsic frame of reference; see Figure 1, right). Intuitively, for some speakers, the version on the left may be more suitable if the relatum is inanimate and a non-possessive construction is used, as in statement (1), and the version on the right would be preferred for a human relatum that is referred to in a possessive construction, as in statement (2). Part of the reason for this intuition is that chairs, unlike humans, arguably do not have very clearly assigned intrinsic left and right sides, which makes the relative reference frame more reliable. In fact, even when the relatum has intrinsic sides, producing and interpreting spatial descriptions dealing with the lateral axis may still incur an increase in processing resources. In their Spatial Framework Theory, Franklin and Tversky (1990) argue that the lateral axis is cognitively challenging due to the lack of salient asymmetries between left and right. In contrast, gravity facilitates the distinction between above and below (vertical axis), and the front and back (sagittal axis) of a body is perceptually and functionally asymmetric.

However, the availability of orientational features alone does not fully account for the systematic preference of a reference frame over another. Languages (and their speakers) deal in different ways with other generic object features such as animacy, as will be discussed in forthcoming sections.
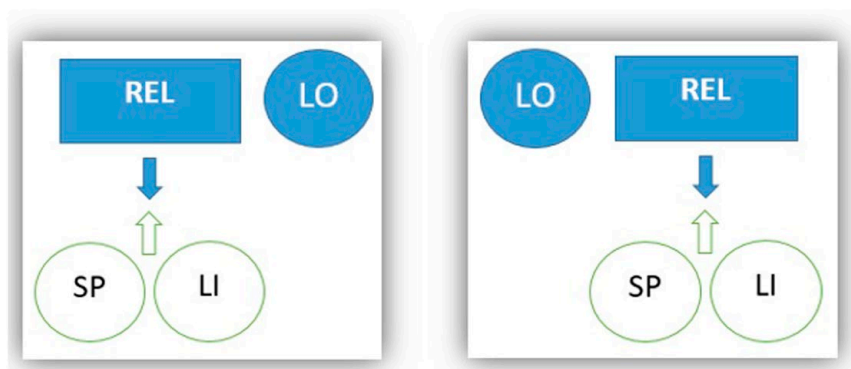
258

Fig. 1. Schematic interpretation of speaker-based perspective choice (relative reference frame; left) and relatum-based perspective choice (intrinsic reference frame; right) for descriptions like (1) and (2). SP = speaker; LI = listener; LO = Locatum; REL = Relatum; grey arrow = relatum's intrinsic direction (front); white arrow = speaker's and listener's view direction.

Furthermore, even though chairs may not have a clearly assigned intrinsic right side, it is still not wrong to refer to *a chair's right*, but the chosen syntactic construction (*to the chair's right* vs. *to the right of the chair*) may play a separate role when choosing a reference frame. The present study aims to clarify what speakers' preferences might be in English and in Spanish. It specifically addresses the impact of animacy and syntactic construction on reference frame selection as potential generic factors that may systematically affect reference frame choices. In the following sections, we will first discuss reference frame choice more generally, and then take a closer look at the two main factors in our study, syntactic construction and animacy.

1.1. SPATIAL PERSPECTIVE CHOICE: IS THERE A DEFAULT FRAME OF REFERENCE?

The literature offers conflicting views as to the existence of a default reference frame in English, and evidence for Spanish is sparse. The earlier literature started out with theoretical considerations based on limited empirical evidence; for instance, Miller and Johnson-Laird (1976) argued that English speakers tend to favour the intrinsic reference frame, and Carroll (1997) extrapolated a similar idea from some empirical findings. In contrast, Levelt (1989) and Levinson (2003) suggested that the speaker's perspective is predominant in English, leading to a preference for the relative reference frame even when the object in question is not directly related to the speaker as relatum. In line with the latter view, Herrmann and Grabowski (1994)

259

argued that listeners should assume that the speaker is using his or her own perspective unless otherwise specified. This is in accordance with studies suggesting that the cognitive effort of taking someone else's perspective is greater than keeping one's own (e.g., Nan, Li, Sun, Wang & Liu, 2016; von Wolff, 2001).

However, based on an increasing body of evidence, it has been repeatedly suggested that perspective choice is highly flexible and context-dependent and may vary relative to different communicative needs (e.g., Schober, 1998; Tenbrink, 2007; Tversky, 1996), such as taking the addressee's perspective to facilitate comprehension (Hund, Haney, & Seanor, 2008; Tversky, 1996). This is in line with the wider literature on different perspectives in discourse (e.g., Dancygier & Sweetser, 2012), which suggests that speakers are highly aware of different viewpoints and adjust their references accordingly. In this light, the idea of a default reference frame may need to be questioned altogether; instead, speakers may flexibly choose from the available repertory according to communicative purposes. Depending on the demands of the situation, they might switch perspectives. This generally happens implicitly, with no explicit signposting in language (Tversky, 1996).

Bowerman (1996) suggested that children born in a particular linguistic and/or cultural context conceptualise space according to the requirements of their native language. This view is consistent with the Whorfian view (Whorf, 1956) that language, to some degree, determines thought (see Danziger, 2011; Levinson, 1996, 2003, for recent advocates of this view as applied to spatial cognition and spatial language). In particular, Danziger (1998) emphasised the need to consider the role that cultural and social factors play in the domain of spatial cognition. However, it may not always be clear whether the tendency to employ a certain reference frame under certain circumstances is due to some specific formal characteristics of the language in question, or to socio-cultural factors that influence individual conceptualisation (Danziger, 1998; Talmy, 2000), or to the situational context itself (Vorwerg & Weiß, 2010). Generally, whenever linguistic constructions are not associated with specific reference frames (e.g., *behind the car* can be interpreted in more than one way), any patterns of preference in speakers of a language must be based on other influencing factors. In some cultures, specific environmental circumstances facilitate the use of absolute reference frames, as in the case of expressions meaning 'downhill' and 'uphill', which speakers ubiquitously use as directions in languages like Tzeltal (Brown & Levinson, 1993) and Gawwada (Tosco, 2012), or directions referring to the north and south banks of a local river, as in Kuuk Thaayorre (Gaby, 2012).

Situational, object-related, or linguistic factors can influence which reference frame the speaker may be employing. Keysar, Barr, and Horton (1998) found that speakers tend to use their own perspective for the production

260

of spatial instructions under time constraints, which suggests that the initial 'instinct' in the utterance-making process is egocentric. Somewhat contrarily, Miller and Johnson-Laird (1976) suggested that interpretation of spatial descriptions depends primarily on the relatum's features: if the object serving as relatum has intrinsic sides, the most likely interpretation is an intrinsic one and vice versa. In our study, we address the impact of object features (beyond the existence of intrinsic sides) by looking at different degrees and aspects of animacy (see Section 1.3), and additionally examine the potential effects of the different linguistic repertories in English and Spanish (see next section).

1.2. SYNTACTIC CONSTRUCTIONS IN ENGLISH AND SPANISH

In English, there are two main ways to describe lateral static configurations: as a possessive construction involving the Saxon genitive (i.e., the *'s* particle denoting possession, as in *X is on Y's left/right*), and in a non-possessive way (*X is to the left/right of Y*). Some authors associate the possessive version primarily with an intrinsic reference frame, and claim that the non-possessive version is more likely to suggest a relative reference frame (Levelt, 1996; Levinson, 2003). Evidence for this view comes from Robinette, Feist, and Kalish (2010), who found that possessive constructions like *the teacup to the teapot's left* triggered an intrinsic interpretation significantly more often than non-possessive constructions such as *the teacup to the left of the teapot*, particularly when relata with an intrinsic front were used.

The motivation for comparing languages in the present study follows up on this linguistic factor. If different constructions lead to the preference of a specific reference frame, the availability of construction types in different languages should affect patterns of reference frame choice. We chose to compare reference frame selection in English with Spanish because of a decisive difference between these languages: Spanish lacks a possessive structure using the Saxon genitive, such as the English *X is on Y's left/right,* to express possession. In Spanish, the most common construction is *X está a la izquierda/derecha de Y* (cf. Romo Simón, 2016), which corresponds to the English non-possessive construction *X is to the left/right of Y*. Alternatively, the speaker may use a marked possessive construction, mainly for clarification in order to refer back to a previously mentioned relatum, as in *Veo Y. X está a su izquierda/derecha* (*I see Y. X is on its left/right*). This construction is superficially similar to the English possessive construction *X is on Y's left/ right*. Nonetheless, it must be noted that these are not equivalent expressions, as the Spanish version is only possible with a possessive adjective that refers back to a previously mentioned relatum, whereas the English construction can stand alone and use any kind of nominal phrase. This difference may prove to be decisive in reference frame choice, since research has suggested

261

that the English possessive construction is often associated with the intrinsic reference frame (Robinette et al., 2010).

Feist and Gentner (2003, p. 394) defined animate objects as those "that are capable of self-determination", acknowledging that the definition may vary cross-linguistically. The role that animacy plays in the construction of different linguistic structures has received considerable interest in linguistic research and its impact has been widely acknowledged in a number of typologically unrelated languages (e.g., Bernárdez, 2016; Yamamoto, 1999). For English, Rosenbach (2002, 2008) studied the relationship between animacy, word order, and grammatical variation concerning the Saxon genitive. Results indicate that animate possessors occur more often in pre-nominal genitive constructions (e.g., *John's house*) than post-nominal genitive constructions (e.g., *the house of John*), whereas the opposite holds for inanimate objects. Since Spanish has no construction equivalent to the Saxon genitive, there cannot be any such effects for this language. Similarly, Feist and Gentner (2003) showed that having an animate relatum (e.g., a hand) supported the use of the preposition *in* rather than *on* to describe the position of the locatum. In Spanish, in contrast, the preposition *en* more or less covers all uses of *in*, *on*, and *at* when these describe spatial relationships (for more extensive information on Spanish prepositions, see López, 1998).

Crucially, a study by Surtees, Noordzij, and Apperly (2012) showed that English speakers from the age of eight onwards tended to consider the intrinsic frame more appropriate in scenes with a human relatum, but considered the relative frame more appropriate for non-human relata. However, their study was only concerned with the sagittal axis (i.e., front/back) and the non-possessive construction. The question thus arises as to whether we can find a similar effect in lateral scenes with different linguistic constructions.

To our knowledge, the impact of animacy on spatial language in Spanish has not been studied. Yet, various kinds of structures are affected by the presence of an animate entity in this language. For example, the preposition *a* (usually translated as *to*) is added to accusative constructions (which mark the direct object of a transitive verb, for example, *him* in the English *Have you seen him?*) when the direct object is a human (Torrego Salcedo, 1999) or an animal, although probably to a lesser extent for the latter. Thus, constructions like *¿Has visto mi monedero?* (*Have you seen my purse?*) require the addition of *a* when the direct object is human, as in *¿Has visto a David?* (*Have you seen David?*), or an animal, as in *¿Has visto al perro?* (*Have you seen the dog?*; *al* results from combining the preposition *a* and

262

the masculine singular definite article *el*). In English, in contrast, the presence of an animate direct object does not trigger any structural changes in accusative constructions.

Thus, animacy plays a role in the choice of syntactic constructions in both languages, albeit in quite dissimilar ways, in areas relevant to spatial cognition and language. This motivates our hypothesis that animacy may affect reference frame selection in the two languages in different ways.

1.4. THE CURRENT STUDY

As outlined in the previous sections, there is evidence that both syntactic construction and animacy may affect reference frame choice in English and Spanish. However, there are still significant gaps. To our knowledge, there are no relevant data on Spanish reference frame choices, little evidence on the actual effects of syntactic construction in English, and even less direct evidence on the effects of animacy. Moreover, in spite of indications that syntactic construction and animacy may be inter-related and interact in their effects on language use, there has been no previous attempt, to our knowledge, to disentangle these two factors. In our study, we address these gaps as follows. In two experiments, we address the impact of animacy on the interpretation of static lateral configurations in English and Spanish when dealing with non-possessive (i.e., *X is to the left/right of Y*) and possessive (i.e., *X is on Y's left/right*) constructions. Along with this, we aim to gather empirical data to address the question of a preferred frame of reference in non-possessive static lateral configurations in English. The reviewed literature motivates the following hypotheses:

1. *Syntactic construction in English:* Based on Levelt's (1996) and Levinson's (2003) claims, supported by Robinette et al.'s (2010) findings on inanimate relata, we hypothesise that English-speaking participants will prefer the relative frame of reference for the non-possessive construction. In line with Robinette et al.'s results, we hypothesise that participants will mainly activate an intrinsic frame of reference for the possessive construction.

2. *Animacy in English:* Similar to results from Surtees et al. (2012) with frontal configurations in English, we expect that relata with a higher animacy level will decrease participants' preference for the relative reference frame.

3. *Syntactic construction in Spanish:* Since Spanish does not have two unmarked syntactic constructions to express attributive possession, we expect syntactic construction to have less of an effect on reference frame selection in Spanish than in English.

263

4.   *Animacy in Spanish:* Animate and human relata in either linguistic construction (i.e., non-possessive or possessive) in Spanish may either (a) lead to a higher preference for the intrinsic reference frame as compared to inanimate relata, or (b) not influence reference frame choice. When presented with scenes with animate relata, Spanish speakers may (c) use the intrinsic frame of reference more often than English speakers, (d) use the relative frame of reference more often than English speakers, or (e) not show a distinctive tendency for either reference frame compared to English-speaking participants.

Although we designed Experiment 1 (English) and Experiment 2 (Spanish) to be sufficiently similar to allow for data comparison across the two languages, we will first report them separately in order to address the impact of animacy and syntactic construction within each language.

2. Experiment 1: English

In Experiment 1, we investigated whether linguistic construction and animacy of the relatum influence reference frame selection in English.

2.1. METHOD

2.1.1. *Participants*

A total of 22 (8 male; mean age = 33.64; *SD* = 13.92) native English speakers with little or no knowledge of Spanish participated in the study. Seven of the participants considered themselves to be fluent in a language other than Spanish. Participants were offered to enter a raffle to win a £30 gift voucher.

2.1.2 *Materials and procedure*

To assess the impact of animacy on the participants' frame of reference choices, we developed an animacy scale based on Rosenbach's (2008, p. 164) scale of inanimate < animate < human. Importantly, the 'inanimate' category was further refined by adding two extra criteria that can easily – although not necessarily – relate to animate entities: sidedness and anthropomorphism. Thus, anthropomorphic inanimate objects were considered more animate than inanimate sided objects, which were in turn considered more animate than inanimate unsided objects. In sum, object types used as relatum were based on the four categorical criteria just mentioned: sidedness, anthropomorphism, animacy, and humanness. Combining these criteria yielded the five different object types shown in (3). We labelled the three inanimate object types as *unsided*, *sided*, and *anthropomorphic* based on the

264

additional criteria mentioned above. The object type labels *animate* and *human* follow Rosenbach. The five object types can be grouped in the following chain from least (unsided) to most (human) human-like:

(3)   Object types used in the current study:
       unsided: –sides, –anthropomorphic, –animate, –human (e.g., a vase)
       sided: +sides, –anthropomorphic, –animate, –human (e.g., a car)
       anthropomorphic: +sides, +anthropomorphic, –animate, –human (e.g., a statue)
       animate: +sides, –anthropomorphic, +animate, –human (e.g., a dog)
       human: +sides, +anthropomorphic, +animate, +human (e.g., a woman)

Each of the five object types comprised six different objects, for a total of 30 objects. All picture stimuli (see examples in Figures 2a and 2b) showed a human avatar facing the front of an object, which served as relatum within the spatial scene. For consistency, all objects shown were photographs. Most objects used as relatum were adapted (i.e., cropped and resized) from freely accessible photos from Wikimedia Commons. The first author photographed the remaining objects. A table listing all the objects used as relatum is included in the 'Appendix'. On both (lateral) sides of the relatum were blue circles representing two balls (A and B), which show the possible locations of the locatum.

Next to the avatar was a speech bubble showing a spatial description using either a non-possessive construction (e.g., *I see a vase. The ball is to the right of the vase*) or a possessive construction (e.g., *I see a vase. The ball is on the vase's right*). While all object types were shown to all participants as a within-subjects factor, linguistic construction was a between-subjects factor with half the participants experiencing only the non-possessive construction (non-possessive condition) and the other half only the possessive constructions (possessive condition). In both conditions, half of the instructions involved the use of *left* and *right*, respectively. Overall, the experiment had a 5 (within-subjects; object type) × 2 (between-subjects; linguistic construction) design.

In addition to the 30 target stimuli scenes, the experiment included 60 filler scenes that used the same type of instruction and linguistic construction as the target scenes, but featured projective terms involving the frontal (e.g., *behind*) and vertical (e.g., *above*) axes. Thus, participants interpreted instructions such as *I see a bucket. The ball is behind the bucket* or *I see a bucket. The ball is on the bucket's back*. Since these instructions were unambiguous in this scenario, they were not included in the analysis.

The experiment was created using OpenSesame 2.9.6 (cf. Mathôt, Schreij, & Theeuwes, 2012). Prior to the actual experiment, participants filled in a questionnaire indicating their age, gender, and knowledge of languages other than English. The main task for participants was to decide whether the
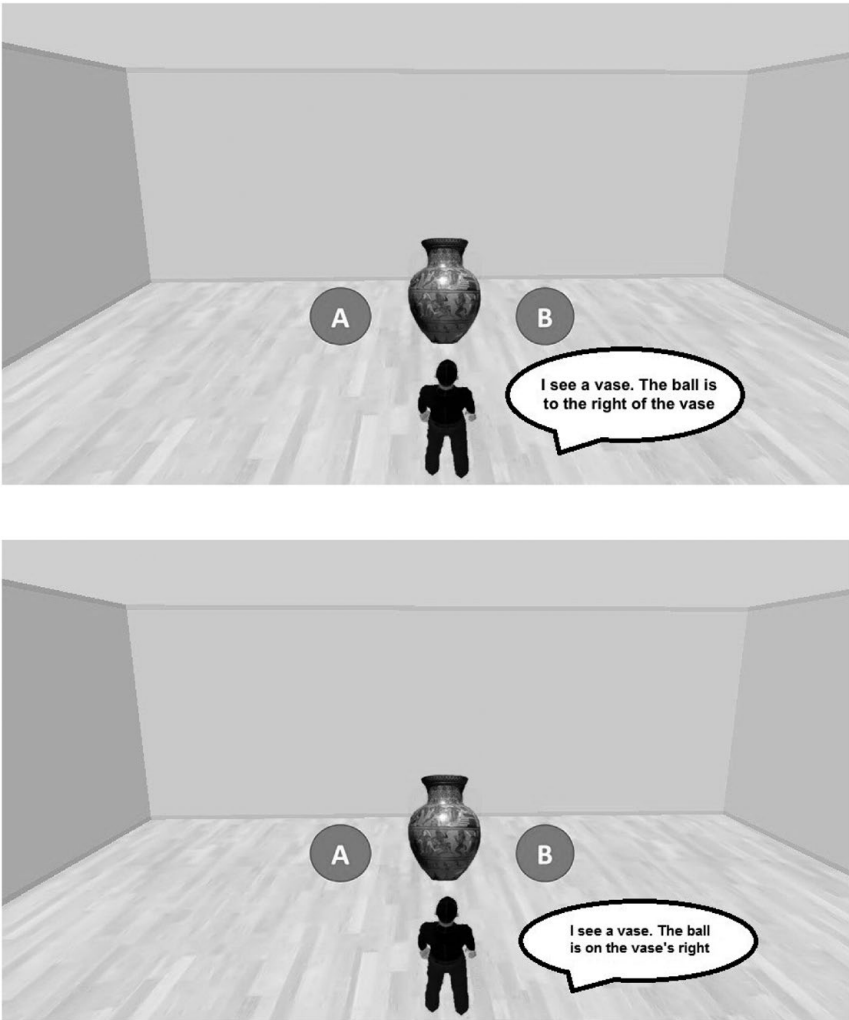
Fig. 2. Example of a target stimulus item presented in Experiment 1 (top: non-possessive condition; bottom: possessive condition).

locatum, i.e., the ball, was in location A or B (see Figure 2) as based on their interpretation of the spatial description presented in the speech bubble. To choose location A, they had to press key A (labelled A) and to choose location B, they had to press key L (labelled B) on the computer's keyboard. To make sure they understood the task, participants received written and spoken instructions and completed one practice trial. Stimuli were presented in three blocks, each containing a set of 30 pictures, for a total of 90 pictures.

266

Each block comprised 10 target (2 per object type) and 20 filler scenes for each participant, in random order within a block. Participants were allowed to take a break between each of the blocks.

The statistical analysis was carried out in R (R Core Team, 2019) using mixed logit models (cf. Baayen, 2008). These models are appropriate for binary response variables (i.e., intrinsic vs. relative frame of reference). Due to the relatively small number of participants in this and the following experiment, we checked whether we had sufficient amounts of observations for all analyses. Specifically, mixed logit models require ten times as many responses or more of the less frequent kind (here either relative or intrinsic frame of reference choices, whichever response occurs less frequently) as there are predictors (i.e., fixed and random factors) in the model (Jaeger, 2011; see Peduzzi, Concato, Kemper, Holford, & Feinstein, 1996, for simulations). Fewer observations of the less frequent kind may lead to overfitting, such that the model would describe the sample and would not allow generalisation to the population. All major analyses presented throughout the paper have sufficient numbers of observations of the less frequent kind (cf. Jaeger, 2011).

The appropriate statistical models were determined through model comparisons (cf. Baayen, 2008). The full model included sentence construction (possessive vs. non-possessive), object type (five levels from unsided to human), and the sentence construction by object type interaction as fixed effects (all centred and sum-coded), and participant and item as random effects. Random slopes for the within-subject factor object type were included for both participant and item (cf. Barr, Levy, Scheepers, & Tily, 2013; Winter & Wieling, 2016). To check if the full fixed and random effects structures were needed, model comparisons were conducted. Fixed and random factors that did not reliably improve model fit were removed from the model. If a model did not converge, the random or fixed effects structure was simplified until the model converged. Data and R scripts for this paper are available at <https://osf.io/krzqd/>.

2.2. RESULTS

We first investigated whether the object type and the sentence construction influenced reference frame choices. Figure 3 shows the relative frequency of intrinsic and relative frames of reference for the five different object types and the sentence construction conditions. Participants in the non-possessive condition overwhelmingly chose the relative frame of reference (305 out of 330 relative responses: 92.42%), whereas participants in the possessive condition overwhelmingly chose the intrinsic frame of reference (318 out of 330 intrinsic responses: 96.36%). In addition, the percentage of intrinsic
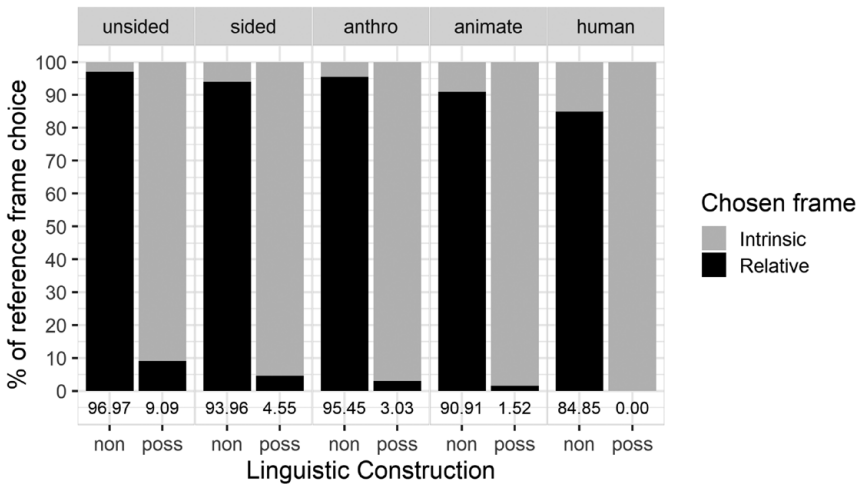
267

Fig. 3. Experiment 1, reference frame choice in English: percentage of responses using a relative vs. intrinsic frame of reference depending on the object type for the non-possessive (non) and possessive (poss) conditions. The numbers below the bars represent percentage of relative frame choices.

responses increases as the degree of animacy rises, suggesting that object type seems to affect the choice of frame of reference, if only to a limited extent.

The final statistical model[1] included sentence construction and object type as fixed effects and no random effects. It showed a significant main effect of both sentence construction (logit estimate = 3.11, std. error = 0.22, z = 14.46, $p < .001$) and object type (logit estimate = 0.72, std. error = 0.2, z = 3.66, $p < .001$) on frame of reference choices. Thus, the possessive construction led to a substantial increase in intrinsic frame of reference choices compared to the non-possessive construction. Frame of reference choices also differed depending on object type. We conducted post-hoc tests using the `emmeans` package in R to determine for which particular object types the frame of reference choices differed reliably. Results only revealed significantly more intrinsic frame of reference choices for human compared to unsided relata (logit estimate = −2.17, std. error = 0.6, z = −3.64, $p < .01$), that is, only for the end points of our animacy continuum.

As the final model includes no random effects, we report the R^2 value for generalised linear mixed effects models (R^2_{GLMM}; Johnson, 2014; Nakagawa & Schielzeth, 2013; Nakagawa, Johnson, & Schielzeth, 2017), which captures the variance explained by a model's fixed factors, to gauge effect size.

---

[1] `glm(RefFrame ~ ConstructionCS+RelatumTypeCS, data = Eng, family = binomial)`

In addition, we report odds ratios (Baguley, 2009). The R^2 value for the final statistical model above is 0.76, suggesting that about three-quarters of the variance in reference frame selections can be explained through the fixed factors sentence construction and object type. Odds ratios were calculated from the final statistical model reported above, but using treatment coding. The odds of choosing the relative frame of reference for the non-possessive construction are 508.45 times larger than for the possessive construction. The odds of choosing a relative frame of reference for unsided relata are 8.77 times larger than for human relata.

2.3. DISCUSSION

In general, the results from Experiment 1 show that reference frame selection in English is affected more by the sentence construction (non-possessive or possessive) that the speaker uses than by the type of object used as relatum. Although there is no one-to-one correspondence between a reference frame and a specific construction, i.e., the reference frame distinction is not grammaticalised as such (Tenbrink, 2007), speakers seem to converge on very strong tendencies. The reason for this may partially lie in the experimental design: linguistic construction was a between-subject factor and participants may have a tendency to be consistent in an experimental setting with respect to their own reference frame choice (Vorwerg, 2009). While increased animacy did lead to an increase in intrinsic reference frame use, this increase was only significant for the endpoints of our animacy continuum.

Overall, our results are in line with our first hypothesis, which stated that participants would prefer a relative reference frame for the non-possessive construction and an intrinsic reference frame for the possessive construction. Thus, the results support Levelt's (1996) and Levinson's (2003) claim that non-possessive constructions involving lateral projective terms typically trigger the use of the relative frame of reference in English, whereas possessive constructions typically trigger the intrinsic frame of reference. This claim had found empirical support in Robinette et al.'s (2010) study, but our findings extend it insofar as we could determine that type of construction affected speakers' choices far more than animacy did. Our results also contradict Miller and Johnson-Laird's (1976) claim that the sidedness of the relatum plays a decisive role in favour of the intrinsic reference frame since we found no significant difference in frame of reference choices for unsided and sided relata.

With respect to the non-possessive construction, Bateman, Hois, Ross, and Tenbrink (2011) suggested that, because of the inherent ambiguity in the construction, co-present interactants would benefit from agreeing on the perspective used. In this regard, our results indicate that listeners'

269

interpretations can be quite systematic, suggesting that disambiguation may not always be needed.

In addition, our results add to those from Surtees et al. (2012), whose study showed that English speakers from the age of eight onwards tended to consider the intrinsic reference frame more appropriate for the non-possessive construction and a human relatum, and the relative reference frame for the non-possessive construction and a non-human relatum. Since their approach only concerned the sagittal axis, the present study does not contradict their findings, but instead suggests that the pattern identified for static frontal configurations does not apply to static lateral ones. This may be related to the idiosyncrasy of the lateral axis and its specific complexity (Franklin & Tversky, 1990).

3. Experiment 2: Spanish

In Experiment 2, we investigate the possible effect of linguistic construction and animacy of the relatum on reference frame selection in Spanish.

3.1. METHOD

3.1.1. *Participants*

A total of 26 native Spanish speakers (19 male; mean age = 48.5; *SD* = 8.39) with little or no knowledge of English participated. One of the 26 participants reported to be fluent in a language other than English (which was not a criterion for exclusion). Two additional participants were excluded, one for misunderstanding the linguistic stimuli and one due to a learning difficulty. As before, participants were offered to enter a raffle to win a €30 gift voucher.

3.1.2. *Materials and procedure*

Experiment 2 employed the same materials and procedure as Experiment 1, except that the linguistic prompt in the speech bubble was presented in Spanish. Again, linguistic construction (possessive vs. non-possessive) was a between-subject factor, and object type (five levels from unsided to human) was a within-subject factor. Again, the visual stimuli showed blue circles on both (lateral) sides of a relatum, which represented two balls (A and B) and indicated the possible locations of the locatum. Participants were asked to locate the ball according to their interpretation of descriptions like *Veo una vasija. La pelota está a la derecha de la vasija* 'I see a vase. The ball is to the right of the vase' in the case of the non-possessive condition, and *Veo una vasija. La pelota está a su derecha* 'I see a vase. The ball is to its right' in the case of the possessive condition.

270

3.2. RESULTS

The data analysis followed the same structure as in Experiment 1. Thus, we first investigated whether object type and sentence construction influenced reference frame choices. Figure 4 shows the relative frequencies of intrinsic and relative reference frame choices for the five different object types and the two sentence construction conditions. The figure shows that participants in both the non-possessive condition and the possessive condition overall preferred the intrinsic over the relative frame of reference (65.90%, i.e., 257 out of 390, intrinsic responses for the non-possessive condition and 93.03%, i.e., 307 out of 330, intrinsic responses for the possessive condition). Unlike the English-speaking participants in Experiment 1, participants in this experiment numerically favoured the relative frame of reference for unsided and sided relata only, but preferred the intrinsic frame of reference for the other object types. Similar to the English-speaking participants in Experiment 1, participants in this experiment overwhelmingly chose the intrinsic frame of reference for the possessive construction.

The statistical analysis procedure for reference frame choices was the same as in Experiment 1. The final statistical model[2] included sentence construction and object type as fixed effects and random slopes of object type for each participant in the random effects structure. The model showed a significant main effect of both sentence construction (logit estimate = 1.73, std. error = 0.6, $z = 2.9$, $p < .01$) and object type (logit estimate = 1.86, std. error = 0.42, $z = 4.42$, $p < .001$) on frame of reference choices. The reliable effect of sentence construction again reflects the fact that the possessive construction led to an increase in intrinsic frame of reference choices compared to the non-possessive construction. The reliable effect of object type shows that frame of reference choices differed depending on object type. Table 1 shows the results from post-hoc tests using the `emmeans` package in R to determine for which particular object types the frame of reference choices differed. The results show that both unsided and sided relata had significantly fewer intrinsic frame of reference choices than anthropomorphic, animate, and human relata.

As the final model includes random intercepts and slopes, we report marginal and conditional R^2 values for generalized linear mixed effects models (R^2_{GLMM}; Johnson, 2014; Nakagawa & Schielzeth, 2013; Nakagawa, Johnson, & Schielzeth, 2017) to gauge effect sizes. As before, we also report odds ratios (Baguley, 2009). The marginal R^2_{GLMM} value for the final statistical model above, which captures the variance explained by the

---

[2] `glmer(RefFrame ~ Construction+RelatumType + (1+RelatumType| Participant), data = Span, family = binomial)`
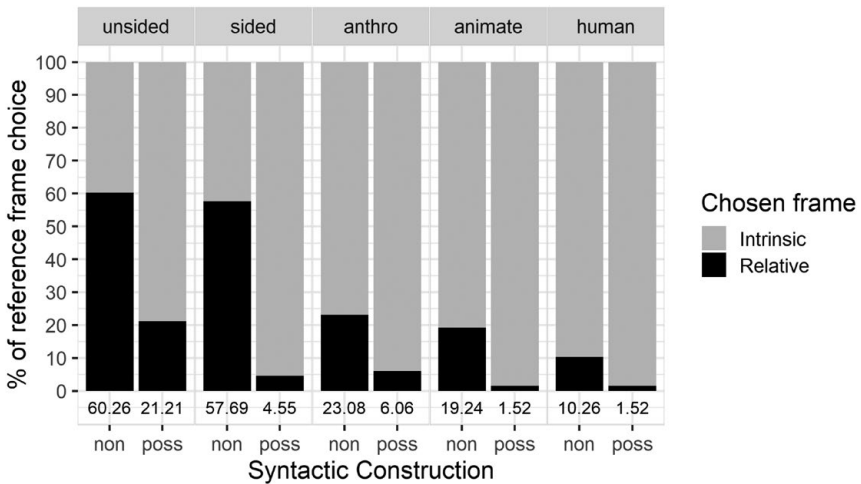
271

Fig. 4. Experiment 2, reference frame choice in Spanish: percentage of responses using a relative vs. intrinsic frame of reference depending on the object type for the non-possessive (non) and possessive (poss) conditions. The numbers below the bars represent percentage of relative frame choices.

model's fixed factors, is 0.35, suggesting that less than half of the variance in reference frame selections can be explained through the fixed factors sentence construction and object type. The conditional R^2_{GLMM} value for the final statistical model above, which captures the variance explained by the model's fixed and random factors, is 0.79, suggesting that the random effects structure contributes about as much to the variance in reference frame selections as do the fixed effects.

As in Experiment 1, we calculated odds ratios using the final statistical model and treatment coding. The odds of choosing the relative frame of reference for the non-possessive construction are 34.27 times larger than for the possessive construction. The odds of choosing the relative frame of reference for unsided relata are 349.39 times larger than for human relata, 50.71 times larger than for animate relata, and 18.95 times larger than for anthropomorphic relata.

3.3. DISCUSSION

The results of Experiment 2 show that both object type and sentence construction affect Spanish native speakers' frame of reference choices. There was an overall preference for the intrinsic frame of reference, which was significantly stronger for the possessive construction than the non-possessive construction. Interestingly, in only two situations did participants

272

TABLE 1. *Statistically significant results from post-hoc tests using the* `emmeans` *package in R*

Comparison	logit estimate	std. error	z value	t value
unsided – anthropomorphic	−2.94	0.77	−3.81	< 0.01
unsided – animate	−3.93	1.12	−3.49	< 0.01
unsided – human	−5.86	1.61	−3.63	< 0.01
sided – anthropomorphic	−2.19	0.59	−3.69	< 0.01
sided – animate	−3.17	0.93	−3.41	< 0.01
sided – human	−5.1	1.43	−3.57	< 0.01

show a numerical preference for the relative frame of reference, namely when the non-possessive construction was used and the relatum was unsided or sided. This is in line with Hypothesis 4a, which stated more relative reference frame choices for inanimate relata compared to animate and human relata as one of the possible outcomes. A direct visual comparison of the Spanish and English results suggests a considerably stronger preference for intrinsic frame of reference choices for Spanish than English. This effect seems to be driven by the non-possessive construction, for which – in contrast to the possessive construction – native Spanish speakers selected an intrinsic frame of reference more frequently than native English speakers. To confirm this, we performed statistical analyses comparing data from the two languages.

3.4. COMPARISON OF EXPERIMENTS 1 AND 2

Our final analysis compares the results from Experiments 1 and 2 in order to address the cross-linguistic questions brought up in Sections 1.2. and 1.3. The experimental design was sufficiently similar for the data to be compared, as the visual prompts (i.e., object types) were identical and the linguistic constructions were as similar as the linguistic repertory of the two languages permits.

The statistical analysis was the same as before, except that Language (English vs. Spanish) was added as a factor to the fixed effects structure. Model comparison for this omnibus analysis was done as described above. The final model[3] revealed a reliable main effect of sentence construction (logit estimate = 2.64, std. error = 0.34, z = 7.77, p < .001) with significantly more intrinsic reference frame choices overall for the possessive compared to the non-possessive construction. There was also a significant main effect of

---

[3] `glmer(RefFrame ~ ConstructionCS + RelatumTypeCS + LanguageCS + ConstructionCS:LanguageCS + RelatumTypeCS:LanguageCS + (1|Partic-ipant), data = EngSpan, family = binomial)`

273

object type (logit estimate = 1.19, std. error = 0.13, $z = 8.91$, $p < .001$), which we will not explore further. Finally, there was a main effect of language (logit estimate = 1.18, std. error = 0.33, $z = 3.63$, $p < .001$) with significantly more intrinsic reference frame choices for Spanish compared to English (the proposed outcome in Hypothesis 4c).

In addition to these main effects, there were significant interactions of sentence construction and language (logit estimate = −1.2, std. error = 0.33, $z = -3.64$, $p < .001$) and object type and language (logit estimate = 0.41, std. error = 0.13, $z = 3.04$, $p < .01$). To explore the sentence construction by language interaction, separate models were fit for the possessive construction and the non-possessive construction. Both models included object type and language as well as their interaction as fixed effects. Model comparison was done as above. Of interest for this section are effects involving the factor language.

The final model for the non-possessive construction[4] showed a main effect of object type (logit estimate = 1.2, std. error = 0.16, $z = 7.55$, $p < .001$) as well as a main effect of language (logit estimate = 2.43, std. error = 0.54, $z = 4.53$, $p < .001$). The latter effect shows that native Spanish speakers selected the intrinsic frame of reference significantly more frequently than native English speakers for the non-possessive construction. In addition, there was a reliable object type by language interaction for the non-possessive construction (logit estimate = 0.55, std. error = 0.16, $z = 3.44$, $p < .001$), just as in the omnibus analysis above.

The final model for the possessive construction[5] showed only a reliable main effect of object type (logit estimate = 1.2, std. error = 0.25, $z = -4.77$, $p < .001$), but included no fixed effects involving language. There were thus similar numbers of relative and intrinsic frame of reference choices across the two languages for the possessive construction. In particular, both native English and native Spanish participants overwhelmingly selected the intrinsic frame of reference for the possessive construction.

The object type by language interaction from the omnibus analysis reflects the fact that animacy affected reference frame selections GRADUALLY in English, with significantly more intrinsic reference frame choices only for human compared to unsided relata (i.e., the endpoints of the animacy continuum), but CATEGORICALLY in Spanish, with significantly more intrinsic reference frame choices for anthropomorphic, animate, and human relata compared to unsided and sided relata.

---

[4] glmer(RefFrame ~ RelatumTypeCS*LanguageCS + (1|Participant), data = non, family = binomial)
[5] glmer(RefFrame ~ RelatumTypeCS + (1|Participant), data = poss, family = binomial)

4. General discussion

Across two experiments, adult participants interpreted spatial descriptions concerning which side (left or right) an object (locatum) was located relative to another object (relatum). Results revealed systematic patterns of reference frame selection, with striking differences between English and Spanish. Although there was a significant object type effect in the two languages, the patterns we see in the post-hoc tests for object type are different. In English, there is a very slight and gradual increase in intrinsic choices as animacy increases, but only the endpoints of this continuum (unsided and human relata) differ significantly from one another. In contrast, in Spanish, there is no gradual increase of intrinsic choices as animacy increases. Instead, there is a categorical distinction such that unsided and sided relata differ reliably from anthropomorphic, animate, and human relata. In addition, the experiments show that the intrinsic frame of reference is predominant when a possessive construction is employed, both in English and in Spanish. However, Spanish speakers choose the intrinsic reference frame more often than English speakers do when a non-possessive construction is used.

Thus, the results open up promising avenues for research on factors guiding reference frame choice. On the one hand, our English data support the claim that choice of grammatical construction can make people think differently about spatial scenes. Specifically, our results show that when different linguistic constructions are available in the linguistic repertoire, these constructions can relate to different reference frames, as Levinson (2003) suggests. On the other hand, our cross-linguistic results highlight the connection between the speakers' mother tongue and spatial cognition, and suggest that analogous constructions (i.e., the non-possessive construction) in different languages can trigger different conceptualisations. In the following, we take a closer look at each of our main results and compare the results for English and Spanish.

4.1. COMPARATIVE ANALYSIS: ENGLISH AND SPANISH

Both languages show very similar patterns regarding the possessive construction, with a clear preference for intrinsic frame of reference choices for all object types. With the non-possessive construction, in contrast, English speakers clearly preferred the relative frame of reference for all object types, while Spanish speakers showed a less clear preference for one reference frame over the other and numerically preferred the intrinsic frame of reference, except for unsided and sided relata. The latter may be related to the concept of a body, as Spanish speakers showed a stronger preference for the intrinsic frame of reference when interpreting non-possessive constructions in static lateral scenes that involved a relatum with a body compared to relata

275

without a body. In contrast, English speakers did not make this distinction, but overwhelmingly interpreted non-possessive constructions to indicate a relative reference frame. Tversky (2005) suggested that bodies constitute a special sort of object within a spatial description because they are experienced both from the inside and from the outside. Bodies are also an essential condition for animacy, since animate entities can typically control their bodies at will under normal circumstances. Therefore, Tversky's suggestion that bodies constitute a special sort of object aligns well with the reference frame choices we found for Spanish, but not for English.

This raises the question of why such a difference is registered in two typologically similar languages. As Talmy (2000) points out, identifying the factors driving reference frame choice is a difficult task, given that employing a certain reference frame might be due to linguistic reasons (i.e., specific formal characteristics of the language) or factors determined by the speaker's environment (cultural, situational, or other). In the following, we argue that it is precisely the interaction of both linguistic and non-linguistic factors that may cause the identified patterns. This is because languages (and their speakers) generally deal with factors such as object properties (which are relevant in specific situations) in different ways.

4.2. LANGUAGE-SPECIFIC DIFFERENCES: THE SYNTACTIC REPERTORY

In studies by Rosenbach (2002, 2008), the use of animate entities was linked with the prenominal genitive construction in English (e.g., *the dog's leg*), which relates to the possessive construction in spatial descriptions. That is, when the idea of possession is applied to an animate possessor, the English language encourages the use of the Saxon genitive. Since Spanish lacks such a construction, we argue, the use of an animate or animate-like object functioning as relatum enables the attribution of 'possessive power' to this object, which – as a corollary – triggers the use of the intrinsic reference frame (possibly as an effect of what is known as inalienability, see Section 4.3). Thus, both English and Spanish are affected by the presence of animate entities in linguistic expressions, including spatial descriptions. The dissimilarities found between English and Spanish partly reside in the fact that the former has two unmarked syntactic alternatives to express attributive possession, whereas the latter has only one (the non-possessive construction). Therefore, the effect of animacy is more salient in Spanish when construing static lateral relationships, because its repertoire encourages the use of one syntactic construction. In English, on the contrary, the availability of two unmarked linguistic alternatives to encode spatial information prevents a salient effect, as animacy typically relates to the possessive construction in that possessive

relations with an animate possessor are more liable to be coded through the Saxon genitive, as Rosenbach pointed out.

4.3. LANGUAGE-SPECIFIC DIFFERENCES: THE IMPACT OF INALIENABLE POSSESSION

The preference in Spanish for an intrinsic interpretation overall and the significantly stronger preference for the intrinsic interpretation for relata with a body (i.e., anthropomorphic, animate, and human) compared to without (i.e., unsided and sided) may be due to a specific notion widely acknowledged in the literature: inalienable possession (Kliffer, 1983; Lamiroy, 2003). This type of possession features an inherent connection between the possessor (the entity that owns another entity) and the possessum (the entity owned by another entity; e.g., Nieuwenhuijsen, 2008), where the possessum is conceived of as being inseparable from the possessor (Heine, 1997). In contrast, alienable possession involves possessor–possessum relationships that are relatively more separable (e.g., a tourist and his or her suitcase). Importantly, inalienable possession may trigger syntactic variations, which differ across languages depending on how much of an impact inalienability has on the language in question. Consider the examples in (4) and (5) from English and Spanish, respectively:

(4)   David lost his leg in an accident
(5)   David perd-ió   la   pierna   en un accidente
      David lose-3PS-PAST the   leg   in an accident
      'David lost his leg in an accident'

   While English requires the use of a possessive marker, Spanish does not. Replacing the definite article with a possessive marker would be grammatical, but marked and redundant in Spanish. In example (5), the possessum *pierna* 'leg' cannot be separated (i.e., alienated) from its possessor (David). As a consequence, *pierna* is preceded by a definite article *la* 'the' instead of the possessive marker *su* 'his/her'. As the part–whole possessive relationship between David and *pierna* is unmistakable, the possessive relationship is conveyed without a possessive marker. Importantly, inalienability does not have the same impact on all languages and in the same way, as what can be considered inalienable varies across languages (Heine, 1997). In particular, the impact of inalienable possession on linguistic constructions appears to be greater in Spanish than in English (Lamiroy, 2003), and overall greater in Romance languages than in Germanic languages (Nieuwenhuijsen, 2008).

   It is worth noting that some elements are more liable to feature an inalienable relationship between possessor and possessum than others. Traditionally, kinship terms and body parts have been analysed as prototypical

277

instances of inalienable possessions (e.g., Barker, 1991; Heine, 1997). This can be explained in terms of conceptual distance, a notion that has been deemed crucial for inalienable possession (Chappell & McGregor, 1989). Thus, conceptually proximal entities are liable to encode inalienable possessive relationships, whereas conceptually distant ones typically encode alienable relations. According to Velázquez-Castillo (1996, p. 36), the conceptual distance between possessor and possessum is partly defined by the "degree of permanency" of the latter. That is, the more permanent a possessum is with respect to its possessor, the more inalienable the relationship is. Since projective terms (e.g., left, front …) typically emanate from body parts, and these have a high degree of permanency, it is not surprising that concepts evoking spatial relations have frequently been considered examples of inalienable possessions.

Of particular relevance is the work by Devylder (2018) on Paamese, an Austronesian language spoken in Vanuatu. Based on empirical research in the field of psychology and perception (e.g., De Vignemont, 2017), Devylder argues that the conceptual distance a possessor perceives between them and a particular body part is smaller for those body parts that they can control and direct. That is, certain body parts, like the limbs or the head, are conceived of as more proximal than others, like internal organs. The distinction is mainly, albeit not exclusively, dependent on the degree of agency of the possessors (humans) over the possessa (their body parts). Importantly, his study shows a correspondence between conceptually proximal body parts and inalienable structures in Paamese, although the author points out that this distinction holds both overtly and/or covertly for many other languages, including English. Again, given that projective terms typically emanate from conceptually proximal body parts, the link between spatial terms and inalienability appears difficult to dispute. In fact, spatial terms have been included on various hierarchies of inalienability (e.g., Chappell & McGregor, 1996; Lichtenberk, Vaid, & Chen, 2011; Nichols, 1992) and, in some languages, they are even more prominent than kin and body parts, as in the case of Mandarin (Chappell & Thompson, 1992) or Ewe (Ameka, 1996).

We suggest that Spanish is another language where inalienability plays a crucial role for encoding spatial scenes. Specifically, animate-like relata may prompt the use of the intrinsic frame of reference in static lateral configurations because the lateral side expressed by the projective term (i.e., *left* or *right*) is understood as an inherent and inalienable element of the relatum when it has animate-like attributes. Hence, both projective terms *izquierda* 'left' and *derecha* 'right' belong to the relatum rather than to the observers. For example, in the spatial description *La pelota está a la izquierda de David* 'The ball is to the left of David', the projective term *left* is conceived of as inherent to the animate relatum, David, and therefore belongs primarily to him, and not to

278

the speaker. Consequently, this spatial description triggers the activation of the intrinsic frame of reference instead of the relative one. The same, we argue, holds for relata in our anthropomorphic and animate categories, since these object types also possess a body. For the non-possessive construction, Spanish speakers show a numerically stronger preference for their own perspective (in a relative reference frame) when the relatum is neither human, animate, nor anthropomorphic, i.e., when the relatum is an entity that is not typically conceived of as something that can possess anything. For example, cars and vases typically do not possess anything. In contrast, there is no such distinction in English because the impact of inalienable possession is not as important as in Spanish.

The differences that we have identified between Spanish and English in this and the previous sections highlight the intricate interplay between the languages we speak and the conceptual patterns we express (such as reference frames). While language, as seen in our study, may not strictly determine conceptual patterns, we can indeed identify strong preferences for a particular reference frame and relate them back to the grammatical resources of the languages, along with animacy. This contributes to the ongoing debates on linguistic relativity, and offers a chance to further explore the degree to which speakers are influenced by their native language.

For instance, the current result opens up an exciting scope for studies exploring reference frame selection in bilingual speakers. Recently, Meakins, Jones, and Algy (2016) found an increase in relative frame choices in speakers of Gurindji who attended tertiary-level education in English. Earlier contributions suggested bilingualism as a possible factor affecting perspective switches in speakers of various languages (e.g., Eggleston, Benedicto, & Balna, 2011, Hernández-Green, Palancar, & Hernández, 2011; Levinson, 2003; Polian & Bohnemeyer, 2011; Romero Méndez, 2011), but did not address this issue directly. However, various authors (e.g., Kleiner, 2004; O'Meara, 2011; Pérez-Báez, 2011) explicitly point to the need for assessing the role of bilingualism in reference frame selection. Studying the effects of this specific discrepancy in Spanish–English bilinguals would thus allow for addressing the question of linguistic relativity from a new angle, as the interplay between linguistic and cognitive aspects is particularly neat in this case.

5. Conclusion

Interpretations of spatial descriptions for lateral static configurations in English and in Spanish are affected by syntactic construction and by animacy, although in different ways. This study sheds light on the question of what factors drive the preference for one reference frame over another in English

279

and Spanish. Based on our results, we propose that the overall preference for the intrinsic frame observed in Spanish in our setting is in large part due to the notion of inalienable possession. Only when the relatum was not a typical or possible possessor, and thus not easily conceived of as an inherent and inalienable part of the relatum, did Spanish speakers tend to abandon their preference for the intrinsic frame of reference and show a significant increase in using their own perspective. In contrast, English speakers selected reference frames primarily on the basis of syntactic construction, suggesting that the grammatical construction made English speakers think differently about spatial scenes. This was perhaps facilitated by the fact that both constructions are unmarked in English, contrasting with Spanish. The concept of inalienable possession does not seem to be as influential in English as it is in Spanish. Instead, if speakers wish to signify a possessive relationship, they can do so by virtue of the possessive construction. Thus, the linguistic features described in the previous section and the differing impact of inalienable possession work together to cause a distinct pattern across the two languages.

Our study hence sheds light on the impact that animacy and construction type might have on spatial interpretations. Further research can complement the present paper by approaching the impact of animacy on static lateral scenes in different languages. Specifically, analyses focusing on either Germanic or Romance languages will serve to enhance the account of the tendencies described in this paper. Finally, future research should also address how Spanish–English bilinguals construe frames of reference in their two languages. Studies of this kind would shed light on the linguistic relativity debate and would provide insight into spatial cognition in bilingual minds.

REFERENCES

Ameka, F. (1996). Body parts in Ewe grammar. In H. Chappell and W. McGregor (eds.), *The grammar of inalienability: a typological perspective on body part terms and the part–whole relation* (pp. 783–840). Berlin: Mouton de Gruyter.

Baayen, R. H. (2008). *Analyzing linguistic data: a practical introduction to statistics*. Cambridge: Cambridge University Press.

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology* **100**(3), 603–617.

Barker, C. (1991). Possessive descriptions. Unpublished doctoral dissertation, UC Santa Cruz. UMI Dissertation Services.

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language* **68**(3), 255–278.

Bateman, J. A., Hois, J., Ross, R. & Tenbrink, T. (2011). A linguistic ontology of space for natural language processing. *Artificial Intelligence* **174**, 1027–1071.

Bernárdez, E. (2016). *Viaje lingüístico por el mundo: Iniciación a la tipología de las lenguas*. Madrid: Alianza Editorial.

Bowerman, M. (1996). Learning how to structure space for language: a crosslinguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (eds.), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.

Brown, P. & Levinson, S. C. (1993). 'Uphill' and 'Downhill' in Tzeltal. *Journal of Linguistic Anthropology* **3**, 46–74.

Carlson, L. A. & Covell, E. (2005). Defining functional features for spatial language. In L. Carlson & E. van der Zee (eds.), *Functional features in language and space: insights from perception, categorization, and development* (pp. 175–190). Oxford: Oxford University Press.

Carroll, M. (1997). Changing place in English and German: language-specific preferences in the conceptualization of spatial relations. In J. Nuyts & E. Pederson (eds.), *Language and conceptualization* (pp. 137–161). Cambridge: Cambridge University Press.

Chappell, H. & McGregor, W. (1989). Alienability, inalienability and nominal classification. In *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society (1989)* (pp. 24–36). Online: doi:10.3765/bls.v15i0.1734.

Chappell, H. & McGregor, W. (1996). Prolegomena to a theory of inalienability. In H. Chappell & W. McGregor (eds.), *The grammar of inalienability: a typological perspective on body part terms and the part–whole relation* (pp. 3–30). Berlin: Mouton de Gruyter.

Chappell, H. & Thompson, S. A. (1992). Semantics and pragmatics of associative *de* in Mandarin discourse. *Cahiers de Linguistique Asie Orientale* **21**(2), 199–229.

Dancygier, B. & Sweetser, E. (eds.) (2012). *Viewpoint in language: a multimodal perspective*. Cambridge: Cambridge University Press.

Danziger, E. (1996). Parts and their counterparts: spatial and social relationships in Mopan Maya. *Journal of the Royal Anthropological Institute* **2**(1), 67–82.

Danziger, E. (1998). Introduction: language, space and culture. *Ethos* **26**(1), 3–6.

Danziger, E. (2011). Distinguishing three-dimensional forms from their mirror-images: Whorfian results from users of intrinsic frames of linguistic reference. *Language Sciences*, **33**, 853–867.

De Vignemont, F. (2017). Agency and bodily ownership: the bodyguard hypothesis. In F. De Vignemont & A. Alsmith (eds.), *The subject's matter: self-consciousness and the body* (pp. 217–237). Cambridge, MA: MIT Press.

Devylder, S. (2018). Diagrammatic iconicity explains asymmetries in Paamese possessive constructions. *Cognitive Linguistics* **29**(2), 313–348.

Eggleston, A., Benedicto, E. & Balna, M. Y. (2011). Spatial frames of reference in Sumu-Mayangna. *Language Sciences* **33**(6),1047–1072.

Feist, M. I. & Gentner, D. (2003). Factors involved in the use of in and on. *Proceedings of the Twenty-Fifth Annual Meeting of the Cognitive Science Society* (pp. 390–395). Online <http://groups.psych.northwestern.edu/gentner/papers/FeistGentner03.pdf>.

Franklin, N. & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General* **119**(1), 63–76.

Gaby, A. (2012). The Thaayorre think of time like they talk of space. *Frontiers in Psychology* **3**, 300. Online: doi:10.3389/fpsyg.2012.00300

Heine, B. (1997). *Possession: cognitive source, forces and grammaticalization*. Cambridge: Cambridge University Press.

Hernández-Green, N., Palancar, E. L. & Hernández, S. (2011). The loanword lado in Otomi spatial descriptions. *Language Sciences* **33**(6), 961–980.

Herrmann, T. & Grabowski, J. (1994). *Sprechen: Psychologie der Sprachproduktion*. Heidelberg: Spektrum.

Hund, A. M., Haney, K. H. & Seanor, B. D. (2008). The role of recipient perspective in giving and following wayfinding directions. *Applied Cognitive Psychology* **22**(7), 896–916.

Jaeger, T. F. (2011). Corpus-based research on language production: information density and reducible subject relatives. In E. M. Bender & J. E. Arnold (eds.), *Language from a cognitive perspective: grammar, usage, and processing. Studies in honor of Thomas Wasow* (pp. 161–198). Stanford, CA: CSLI Publications.

281

Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution* **5**(9), 944–946.

Keysar, B., Barr, D. J. & Horton, W. S. (1998). The egocentric basis of language use: insights from a processing approach. *Current Directions in Psychological Science*, **7**(2), 46–50.

Kleiner, L. F. (2004). Review of the book *Space in Language and Cognition: Explorations in Cognitive Diversity* by S. Levinson. *Journal of Pragmatics* **36**, 2089–2099.

Kliffer, M. D. (1983). Beyond syntax: Spanish inalienable possession. *Linguistics* **21**, 759–794.

Lamiroy, B. (2003). Grammaticalisation and external possessor structures in Romance and Germanic languages. In M. Coene & Y. D'Hulst (eds.), *From NP to DP. Volume II: the expression of possession in noun phrases* (pp. 257–280). Amsterdam: John Benjamins.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA.: MIT Press.

Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial descriptions. In P. Bloom, M. A. Peterson, L. Nadel & M. Garret (eds.), *Language and space* (pp. 77–108). Cambridge, MA: MIT Press.

Levinson, S. C. (1996). Frames of reference and Molyneux's Question: crosslinguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel & M. l Garret (eds.), *Language and space* (pp. 463–492). Cambridge, MA: MIT Press.

Levinson, S. C. (2003). *Space in language and cognition*. Cambridge: Cambridge University Press.

Lichtenberk, F., Vaid, J. & Chen, H. (2011). On the interpretation of alienable vs. inalienable possession: a psycholinguistic investigation. *Cognitive Linguistics* **22**(4), 659–689.

López, G. , A. (1998). *Gramática del español. III. Las partes de la oración*. Madrid: Arco/Libros.

Mathôt, S., Schreij, D. & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* **44**(2), 314–324.

Meakins, F., Jones, C. & Algy, C. (2016). Bilingualism, language shift and the corresponding expansion of spatial cognitive systems. *Language Sciences* **54**, 1–13.

Miller, G. A. & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Nakagawa, S., Johnson, P. C. & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface* **14**. Online: doi.org/10.1098/rsif.2017.0213

Nakagawa, S. & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* **4**(2), 133–142.

Nan, W., Li, Q., Sun, Y., Wang, H. & Liu, X. (2016). Conflict processing among multiple frames of reference. *PsyCh Journal* **5**, 256–262.

Nichols, J. (1992). *Linguistic diversity in space and time*. Chicago, IL: University of Chicago Press.

Nieuwenhuijsen, D. (2008). La posesión inalienable en español y su traducción en varias lenguas germánicas y románicas: una comparación. *Hermēneus. Revista de Traducción e Interpretación* **10**, 1–19.

O'Meara, C. (2011). Spatial frames of reference in Seri. *Language Sciences* **33**, 1025–1046.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology* **49**(12), 1373–1378.

Pérez-Báez, G. (2011). Spatial frames of reference preferences in Juchitán Zapotec. *Language Sciences* **33**, 943–960.

Polian, G. & Bohnemeyer, J., (2011). Uniformity and variation in Tseltal reference frame use. *Language Sciences* **33**(6), 868–891.

R Core Team (2019). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online <https://www.R-project.org/>.

Robinette, L. E., Feist, M. I. & Kalish, M. L. (2010). Framed: factors influencing reference frame choice in tabletop space. In S. Ohlsson & R. Catrambone (eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 1064–1069). Austin, TX: Cognitive Science Society.

282

Romero Méndez, R., (2011). Spatial frames of reference and topological descriptions in Ayutla Mixe. *Language Sciences* **33**(6), 915–942.

Romo Simón, F. (2016). Un estudio cognitivista de las preposiciones espaciales del español y su aplicación a la enseñanza de E/LE. Unpublished doctoral dissertation, Universitat Autònoma de Barcelona. Retrieved from <https://ddd.uab.cat/pub/tesis/2016/hdl_10803_384719/frs1de1.pdf>.

Rosenbach, A. (2002). *Genitive variation in English: conceptual factors in synchronic and diachronic studies*. Berlin/New York: Mouton de Gruyter.

Rosenbach, A. (2008). Animacy and gramatical variation: findings from English genitive variation. *Lingua* **118**, 151–171.

Schober, M. F. (1998). Different kinds of conversational perspective-taking. In S. R. Fussell & R. J. Kreuz (eds.), *Social and cognitive psychological approaches to interpersonal communication* (pp. 145–174). Mahwah, NJ: Lawrence Erlbaum.

Surtees, A. D. R., Noordzij, M. L. & Apperly, I. A. (2012). Sometimes losing your self in space: children's and adults' spontaneous use of multiple spatial reference frames. *Developmental Psychology* **48**, 185–191.

Talmy, L. (2000). *Toward a cognitive semantics (Vol 1, concept structuring systems)*. Cambridge, MA: MIT Press.

Tenbrink, T. (2007). *Space, time, and the use of language*. Berlin: Mouton de Gruyter.

Tenbrink, T. (2011). Reference frames of space and time in language. *Journal of Pragmatics* **43**, 704–722.

Torrego Salcedo, E. (1999). El Complemento Directo Preposicional. In I. Bosque & V. Demonte (eds.), *Gramática Descriptiva de la Lengua Española (Vol. 2: Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales)* (pp. 1779–1806). Madrid: Espasa-Calpe.

Tosco, M. (2012). The grammar of space of Gawwada. In M. Brenzinger & A.-M. Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics, Cologne (August 2009)* (pp. 523–532). Cologne: Rüdiger Köppe.

Tversky, B. (1996). Spatial perspective in descriptions. In P. Bloom, M. A. Peterson, L. Nadel & M. Garret (eds.), *Language and space* (pp. 463–492). Cambridge, MA: MIT Press.

Tversky, B. (2005). Form and function. In L. Carlson & E. van der Zee (eds.), *Functional features in language and space: insights from perception, categorization, and development* (pp. 331–348). Oxford: Oxford University Press.

Velázquez-Castillo, M. (1996). *The grammar of possession: inalienability, incorporation and possessor ascension in Guaraní*. Amsterdam: John Benjamins.

von Wolff, A. (2001). *Transformation und Inspektion mentaler Umraumrepräsentationen: Modell und Empirie*. Vienna: GeoInfo Series.

Vorwerg, C. (2009). Consistency in successive spatial utterances. In K. Coventry, T. Tenbrink, & J. Bateman (eds.), *Spatial language and dialogue* (pp. 40–55). Oxford: Oxford University Press.

Vorwerg, C. & Weiß, P. (2010). Verb semantics affects the interpretation of spatial prepositions. *Spatial Cognition and Computation* **10**, 247–291.

Whorf, B. L. (1956). *Language, thought, and reality*. New York: Technology Press of MIT and Wiley.

Winter, B. & Wieling, M. (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. *Journal of Language Evolution* **1**(1)*,* 7–18.

Yamamoto, M. (1999). *Animacy and reference: a cognitive approach to corpus linguistics*. Amsterdam: John Benjamins.

Zlatev, J. (2007). Spatial semantics. In D. Geeraerts & H. Cuyckens (eds.), *The Oxford handbook of cognitive linguistics* (pp. 318–350). Oxford: Oxford University Press.

Appendix

List of objects used as relatum in target scenes

Object type	Relatum	Target side indicated by the stimulus
Unsided	Tree	Left
	Rock	Left
	Table	Left
	Bottle	Right
	Barrel	Right
	Vase	Right
Sided	Tractor	Left
	TV	Left
	Car	Left
	Motorbike	Right
	Chair	Right
	Bike	Right
anthropomorphic	Robot	Left
	Gnome	Left
	Sculpture	Left
	Scarecrow	Right
	Mannequin	Right
	Statue	Right
animate	Sheep	Left
	Cow	Left
	Eagle	Left
	Gorilla	Right
	Cobra	Right
	Dog	Right
human	Man 1 (Daniel)	Left
	Woman 1 (Emma)	Left
	Man 2 (David)	Left
	Woman 2 (Julia)	Right
	Man 3 (Samuel)	Right
	Woman 3 (Laura)	Right

284