Language and Cognition **13** (2021), 503–533. doi:10.1017/langcog.2021.11
© The Author(s), 2021. Published by Cambridge University Press

The semantics of spatial demonstratives in Spanish: a Demonstrative Choice Task study

EMANUELA TODISCO*

Universitat de les Illes Balears, Departament de Filologia Espanyola, Moderna i Clàssica

ROBERTA ROCCA

The University of Texas at Austin, Department of Psychology and Aarhus Universitet, Interacting Minds Centre

AND

MIKKEL WALLENTIN

Aarhus Universitet, Department of Linguistics, Cognitive Science and Semiotics, Aarhus Universitet, Interacting Minds Centre, and Aarhus Universitet, Centre of Functionally Integrative Neuroscience

(*Received* 23 *December* 2020 – *Revised* 11 *May* 2021 – *Accepted* 14 *May* 2021 – *First published online* 15 *July* 2021)

ABSTRACT

Demonstratives (*this*/*that* in English) are pivotal in communication. In this study, we show that semantic features of referents systematically influence speakers' choices of demonstrative forms for Spanish nouns in the absence of a guiding context. We used the Demonstrative Choice Task (DCT), previously applied to two-term demonstrative languages (Danish, English, and Italian), and applied it to Spanish, a three-term demonstrative system (*este*/*ese*/*aquel*), to test if the semantic dimensions driving demonstrative choice overlap with those found for English. 1,639 native Spanish speakers were presented with 480 nouns rated along 76 semantic features and were asked to match each noun with a demonstrative. We found that demonstratives are influenced by the same semantic factors as two-term languages, such as manipulability, valence, and the self. In Spanish, these semantic factors predict the demonstrative choice between the proximal *este*/*a* and a combination of medial and distal forms *ese*/*a* and *aquel*/*la*. Additional semantic factors affect speakers' preferences for *ese*/*a* versus *aquel*/*a* (e.g., visuality and time). We conclude that many

[*] Address for correspondence: Emanuela Todisco. e-mail: e.todisco@uib.es

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

of the semantic attractors influencing the choice of demonstratives are constant across languages, independent of the number of terms characterizing the demonstrative system, and provide a window into the landscape of meaning subserving linguistic reference.

KEYWORDS: Spanish spatial demonstratives, semantics, DCT, cross-linguistic invariants

1. Introduction

1.1. SPATIAL DEMONSTRATIVES AS DEICTIC EXPRESSIONS

Demonstratives (*this*/*that* in English) are cross-linguistically widespread deictic expressions. They are crucial for conveying information about space, time, person, and discourse in exophoric and endophoric contexts (Diessel, 1999, 2005; Diessel & Coventry, 2020), as well as for manipulating the interlocutor's focus of attention during interaction (Kita, 2003; Rocca, Wallentin, Vesper, & Tylén, 2019b; Todisco, Guijarro-Fuentes, Collier, & Coventry, 2020).

Demonstratives are among the most frequent and FLEXIBLE words in the lexicon (Diessel, 2013, 2014; Leech, Rayson, & Wilson, 2014). In different situations, expressions such as 'This one' may refer to an object, a person, an event, or a thought. The advantage of such flexibility is that demonstratives can be used in a variety of linguistic and pragmatic contexts. The disadvantage that follows is that uttering a demonstrative will not always ensure identifiability of the intended referent, as the following examples show. In the sentence 'This is what I think about you: you are a lovely person', the cataphoric use of the pronoun *this* is disambiguated and easily identifiable by the sentence that immediately follows (i.e., you are a lovely person). Nevertheless, in the sentence 'Give me that thing', the adjectival use of *that* is not easily identifiable without additional cues. These cues are usually provided by nonverbal components of the context of utterance, including pointing gestures and eye-gaze trajectories (Kita, 2003; Talmy, 2018; Todisco et al., 2020).

Combining verbal and nonverbal referencing strategies, that is, pointing and gazing at a specific referent while using a demonstrative, facilitates the task of disambiguating the referent, compared to contexts where the speaker only uses one modality, either verbal or nonverbal (Cooperrider, 2016; Fricke, 2014; Iverson, Tencer, Lany, & Goldin-Meadow, 2000; Kita, 2003; Todisco et al., 2020). Talmy (1983, 1988) and Kemmerer (2006, p. 1608) highlighted that demonstratives alone do not encode precise degrees of remoteness from the speaker (or DEICTIC CENTRE). Sentences like 'Here in this room' and 'Here in this galaxy' are both perfectly grammatical, despite the huge differences in scale, because the semantic space demonstratives carve is pragmatically modulated by the combination of nonverbal cues, communicative intention of the

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

speaker, physical position, and salience of referents in the contextual communicative scenario (see Diessel & Coventry, 2020; Kemmerer, 2006, p. 1608). Moreover, the choice of one specific demonstrative rather than another also conveys SOME disambiguating information about the intended referent. In the utterance, 'I know *that* guy', for example, the use of *that* – rather than the perfectly grammatical alternative *this* – hints at the referent being located at some physical or social distance from the speaker (Peeters, Krahmer, & Maes, 2021; Rocca, Tylén, & Wallentin, 2019a; Rocca & Wallentin, 2020).

In the present study, we aim at clarifying the role of the SEMANTICS of referents in the choice of demonstrative forms used to indicate them, using a simple and versatile experimental paradigm called the Demonstrative Choice Task (henceforth DCT). Previous studies have used this paradigm to investigate semantic factors driving demonstrative choice in English (a two-term demonstrative system with *this*/*these* as proximal forms and *that*/*those* as distal forms; Rocca & Wallentin, 2020) and other two-term European languages (Danish and Italian, though only a small number of semantic dimensions were investigated for these languages; Rocca & Wallentin, 2019a In this study, we use the DCT to delve into the semantics of demonstrative choice in Spanish, a language with three demonstrative forms (*este* 'this' / *ese* 'this–that' / *aquel* 'that'). Comparing patterns of demonstrative use in the DCT for Spanish with findings for a two-term language, such as English (*this*/*that*) will contribute to shedding light on whether similar clusters of semantic dimensions drive demonstrative choices across languages, EVEN WHEN LANGUAGES DIFFER IN THE NUMBER OF TERMS AVAILABLE IN THEIR DEMONSTRATIVE SYSTEMS. Furthermore, results from this study will allow us to gather insights on whether the use of medial demonstrative forms (i.e., *ese*) versus distal ones (i.e., *aquel*) can be explained by specific semantic dimensions of the referent, or whether the two are entirely redundant from a semantic perspective.

1.2. PERCEPTUAL VS. PSYCHOLOGICAL FACTORS INFLUENCING DEMONSTRATIVE CHOICE

The speaker's choice of a specific demonstrative form results from a complex process of resolving the mappings between a number of multimodal perceptual, attentional, and pragmatic cues in the minds of both the speaker and the hearer and their relations to the intended referent. Cues functional to identifying the referent can largely be categorized as PERCEPTUAL or PSYCHOLOGICAL factors. PERCEPTUAL cues are those pieces of information contributed, for example, by the physical circumstances of the utterance. Distance between the referent and the speaker is perhaps the most intuitive one, and the most widely studied and debated. Distance modulates the choice

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

of demonstrative based on a near/far dichotomy (Di Pellegrino & Làdavas, 2015), such that the chosen form encodes information on whether the referent is proximal or distal from the speaker (or deictic centre). Many studies have pointed to space and distance between speaker and referent as being the most prominent parameter determining speakers' choice of a specific demonstrative form (Caldano & Coventry, 2019; Coventry, Valdés, Castillo, & Guijarro-Fuentes, 2008; Diessel, 1999, 2005; Hunley & Lourenco, 2018).

In the context of these studies, DISTANCE is often understood as a physical space separating the speaker and the referent. However, recent work by Bocale and Cologna (2020) has pointed out that distance in communication should be better understood as incorporating a variety of subjective factors. An adequate notion of distance does not merely resolve in an indicator of physical space, but is a complex notion which incorporates both ABSOLUTE and PERCEPTIONAL/PERSPECTIVAL distance relationships between interlocutors. ABSOLUTE DISTANCE in this terminology is the space between two discourse elements, such as the speaker and the intended referent. PERCEPTIONAL/PERSPECTIVAL DISTANCE includes a displaced reference point, such as the perspective of the hearer involved in the communicative interaction (see Rocca et al., 2019b).

Distance is thus a prominent but not the only perceptual factor affecting the choice of demonstratives. It has been pointed out, for example, that visibility of the referent may also systematically influence demonstrative choice, as shown by Coventry and colleagues in a study using an adapted version of the psycholinguistic elicitation paradigm known as the Memory Game for Object Location (Coventry, Griffiths, & Hamilton, 2014). In this study, adult English monolingual speakers were asked to point to and verbally describe the position of discs placed at varying distances, with the visibility of discs being manipulated by covering them with either an opaque or a clear container. Results showed that the proximal demonstrative form *this* was used less often when the objects were covered with the opaque container, meaning that whether the object is visible or not also affects demonstratives choice and usage.

Alongside perceptual cues provided by the context of the utterance, the mapping between demonstratives and referents is also mediated by PSYCHOLOGICAL factors, which depend on more abstract, cognitive processes, such as memory and emotions, rather than on perceptual ones. Social proximity between speaker and hearer, and the perspective of the hearer (Küntay & Özyürek, 2002; Peeters, Hagoort, & Özyürek, 2015; Peeters et al., 2021), for example, can determine whether the speaker chooses a proximal, medial, or a distal demonstrative form. Knowledge of the functional status of the referent (i.e., its affordance for manual grasp, see Coventry et al., 2008; Coventry et al., 2014; Rocca et al., 2019a), as well as whether or not the referent is within or

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

outside attentional focus (Strauss, 2002), are other psychological factors explored in the literature.

Evidence for the role of a variety of perceptual and psychological factors in speakers' choices of demonstrative forms (and, therefore, in reference resolution on the addressee's side) does not necessarily entail that ALL of these parameters do ALWAYS (or always EQUALLY) play a role in speakers' demonstrative choices or in the identification of the intended referent. A strict one-to-one correspondence, e.g., one specific parameter fully describing demonstrative choice in a specific language, is reductive and may lead to a misinterpretation of how the identification of the intended referent works. Parameters are not mutually exclusive; a RANGE of potentially relevant parameters is available to the speaker and hearer, from which they SELECT the most salient ones in accordance with the information they want to convey or with communicative needs/constraints (Peeters et al., 2021; Todisco, Guijarro-Fuentes, Gudde, & Coventry, unpublished observations). This combination of systematicity and flexibility is what makes demonstratives such an efficient, pervasive, and cross-linguistically widespread communication tool: all languages have SOME demonstrative forms, although the number of demonstrative terms and parameters affecting demonstrative choice can vary (Diessel, 2005, 2013; Diessel & Coventry, 2020; Levinson, 2018). Providing a full description of which factors drive demonstrative choice and use, how their importance is modulated by context, and the extent to which these dynamics generalize across languages thus provides a key contribution to unpacking the pivotal mechanisms underlying the complex functioning of linguistic reference in interaction (Peeters et al., 2021).

In this study, we delved into a fundamental conceptual and psychological factor of demonstratives: their semantics. The relationship between semantics and deixis has been extensively discussed in the previous literature, but, as highlighted by Dufter (2015), a great part of these studies focuses on the interface between semantics and pragmatics in discourse (e.g., the function of deictic expressions in discourse, or the role of a speaker's intentions in producing a deictic reference), and they mostly use qualitative tools. However, semantics and demonstrative reference are interwoven at a more fundamental level. The choice of demonstrative forms may in fact be constrained by which entities they are meant to refer to, especially when no other additional contextual cues, such as addressee position or physical distance, are provided. The semantics of such entities (and, thus, of the nouns denoting them) influence which demonstrative form are used, and, in turn, the use of specific demonstrative forms and/or their position provides information about the referring noun and/or the intended object (Jungbluth & Da Milano, 2015).

Previous studies on the semantics of demonstratives have mostly focused on a few selected, hypothesis-driven semantic factors at a time, or as embedded in

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

interaction and physical space. Paradigms like the David Wilkins Demonstrative Questionnaire (DWDQ, 1999), the Memory Game for Object Location (Coventry et al., 2014), and conversation data collection (Küntay & Özyürek, 2002) are example of elicitation methods used to investigate factors driving demonstrative choice (see also Levinson, Cutfield, Dunn, Enfield, & Meira, 2018, for a cross-linguistic detailed review). The DWDQ manipulates 25 parameters – among which are: distance, interlocutors' focus of attention, visibility, possession, and mention of the referent during conversation – while assessing one factor at a time (Lin, 2013; Wilkins, 1999). Coventry and collaborators (2014) assessed that ownership, during the Memory Game paradigm, was affecting demonstrative choice and resulting in a larger use of *this*.

Thus, while perceptual factors and single psychological factors have been widely explored and debated, and differences in their role across languages extensively discussed (Levinson et al., 2018), the role of semantics in demonstratives choice has only recently started to receive attention in the experimental literature (Rocca et al., 2018, 2019a, 2019b; Rocca & Wallentin, 2020). The first attempt to provide a comprehensive overview of the semantic dimensions on demonstrative choice which both isolates the contribution of semantics from perceptual and pragmatic factors and probes the role of a large set of psychologically relevant dimensions was provided by the online questionnaire paradigm DCT (Rocca et al., 2018, 2019a, 2019b; Rocca & Wallentin, 2020).

1.3. THE DEMONSTRATIVE CHOICE TASK: MAPPING THE ROLE OF SEMANTICS

Rocca and collaborators (Rocca et al., 2018, 2019a, 2019b; Rocca & Wallentin, 2020) studied the effect of semantic features of concrete referents/nouns in demonstrative choice for two-term demonstrative systems. By asking participants to match nouns and demonstrative forms without priming them with any contextual cues, they showed that speakers of different European languages (English, Danish, and Italian) converge on using proximal demonstratives when referring to words denoting harmless, familiar objects (like apples), and distal demonstratives for harmful, threatening objects (such as tigers). Moreover, Rocca and collaborators (2019a) found that manipulability of the intended referent, a semantic dimension combining both perceptual (size), psychological (harmfulness), and ontological (animacy) dimensions of the intended referent, influences the choice of a proximal demonstrative (i.e., *this*) over a distal form (i.e., *that*) in two-term demonstrative systems, such as English, Italian, and Danish. In a subsequent study including a larger number of both words, semantic dimensions, and participants (Rocca & Wallentin, 2020), it was found that demonstrative choice of the English proximal form (i.e., *this*) over a distal form (i.e., *that*) was related to a set of

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

additional semantic factors, including valence, arousal, loudness, motion, time, and more generally the self as a multidimensional perceptual and psychological construct.

The results reported by Rocca and collaborators suggest that demonstrative choice depends on the interaction of a number of semantic features of a referent relative both to the speaker's physical AND psychological spaces. More specifically, objects can be closer or further away in space, but also in time (e.g., *this* time), or phenomenologically and emotionally (e.g., Don't be *that* guy) (Peeters et al., 2021; Rocca & Wallentin, 2020). Distance is evaluated not only relative to the physical anchoring of the speaker in the immediate perceptual space, but also, in what Bühler (1934/2011) and Stukenbrock (2014, 2015) define as Deixis am Phantasma, relative to the position of the speaker in an imagined space. The dynamic deictic displacements have been described by Stukenbrock (2014) as a CONTINUOUS SHIFTING OF ORIGOS among physical and psychological reference spaces and dimensions. When no physical distance information can be or needs to be conveyed, the reference space becomes an imaginary one, organized around dimensions such as familiarity and affection (e.g., sociocentric parameter in demonstrative choice, Peeters et al., 2015, 2021). In line with this, the DCT results suggest that the imagined space is organized around the self and its needs and motivations. The notion of self conflates the function of physical, bodily, and psychological reference spaces into one multidimensional hyperspace, and the referent's location within this space is what determines the choice of one demonstrative over another.

While the results from large-scale DCT studies in English (Rocca et al., 2019a; Rocca & Wallentin, 2020) are reliably consistent across participants, no evidence exists yet that these patterns generalize to the dynamics of demonstrative use across languages which differ in their demonstrative inventory. One of the strengths of the DCT paradigm, however, lies in the possibility of seamlessly adapting it to different languages, regardless of the number of demonstratives a specific system has. Showing that semantic patterns replicate not only in languages with two-term demonstrative systems, but also across languages with more than two demonstrative terms, would provide a strong argument supporting the hypothesis that these mechanisms capture a fundamental aspect of the relation between human cognition, semantics, and linguistic reference. This is the rationale behind the present study, where we extended the DCT study by Rocca and Wallentin (2020) to Spanish, in order to both probe the replicability of the results obtained so far, and to characterize the role of additional demonstrative forms, such as medial forms in the Spanish three-term system, in the interplay between demonstrative reference and the self.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

1.4. DEMONSTRATIVE TERMS IN SPANISH

Spanish has a three-term demonstrative system which consists of a proximal form, *este/a* 'this', a medial form, *ese/a* 'this/that', and a distal term, *aquel/la* 'that'. The three terms present masculine and feminine, as well as singular and plural, forms which convey gender and number information in accordance with the noun they accompany (adnominal demonstratives: e.g., *esta luna* 'this moon') or substitute (e.g., *estos* 'these', referring to the plural referents 'books'). The proximal–medial–distal distinction has been suggested to designate increasingly remote concentric circles around the speaker (Hottenroth, 1982, p. 133). Traditionally, the adnominal/pronominal proximal form *este/a* 'this' has been said to identify a portion of space close to the speaker; the medial form *ese/a* 'this/that' identifies a medial distance from the speaker (or hearer), and *aquel/la* 'that' is used for portions of space located far away from the speaker (and hearer). This division is partly supported by experimental evidence (Coventry et al., 2008). Using the Memory Game for Object Location (Gudde, Griffiths, & Coventry, 2018), Coventry and collaborators (2008) found that speakers of Spanish also tend to carve up space along a peri-personal (within reach) versus extra-personal (outside reach) distinction. Participants used more instances of the proximal demonstrative *este/a* 'this' in the peri-personal space compared to extra-personal space. Interestingly, however, Coventry and colleagues did not look at the distinction between usage patterns of the medial (*ese/a*) and the distal (*aquel/la*) demonstrative in the extra-personal space, which therefore remains largely unexplored. Moreover, in the majority of the above-mentioned studies, semantic distinctions between referents are 'controlled for' or neutralized rather than thematized as a factor of interest (Coventry et al., 2008; Jungbluth, 2003). To our knowledge, there is no empirical study systematically and comprehensively describing the use of Spanish demonstrative forms as a function of semantic features of the referent when no contextual cue is provided and without the interaction with an addressee.

1.5. THE PRESENT STUDY

In the present study, we used the DCT to investigate the relationship between semantics and demonstrative choices in Spanish. The aims of this study were twofold. First, we aimed at investigating whether and how semantic features influence demonstrative choices in a language with a three-term demonstrative system. We expected to replicate results from the English two-term system, with semantic features related to referent manipulability, valence, and to a psychological notion of SELF driving preference for proximal vs. other forms (cf. Rocca & Wallentin, 2020). To this end, Spanish data were collected for the

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

same set of words used in this study. These words comprised a set of words, previously given comprehensive semantic annotations by Binder and collaborators (2016). Second, we aimed at exploring usage patterns and potential semantic specificities of medial and distal forms (i.e., *ese/a* and *aquel/la* in Spanish) in contexts where fine-grained spatial distinctions are not needed. With respect to the latter, two possible scenarios could be expected. In contexts where conveying precise information on the referent's physical location is not necessary, the distinction between medial and distal forms could:

(a) *disappear, leading to a collapse onto a binary near/far opposition*. If this were the case, we would expect to see either a random distribution of the two non-proximal demonstratives, or a diminished use of one of the two non-proximal demonstratives, most probably the distal form *aquel/la* 'that' in favour of the medial form *ese/a* 'this-that', in line with a trend of reduction already identified in Argentinian Spanish (Pérez-Saldanya, 2015) and the Majorcan Catalan (Todisco, Guijarro-Fuentes, & Coventry, 2021). The Argentinian Spanish three-term system (*este* 'this' / *ese* 'this-that' / *aquel* 'that') is in fact characterized by the co-existence of the two forms *ese/a* and *aquel/la* for the same dimension (the medial/distal space), where the distal demonstrative *aquel/la* is considered more formal than the medial demonstrative *ese/a* (Pérez-Saldanya, 2015, p. 116), but semantically equivalent. Majorcan Catalan demonstratives system (*aquest/aqueix/aquell*) presents a reduction to two-terms (*aquest/aquell*) when information about physical distance is conveyed (Todisco et al., 2021).

(b) *adapt to specific semantic factors and dimensions*. The choice of the medial *ese/a* versus the distal term *aquel/la* could be made on the basis of semantic factors. In other words, speakers may use a multidimensional mapping between demonstratives and semantics, such that particular semantic dimensions map onto the medial/distal gradient of *ese/a-aquel/la*, while other dimensions may map onto a proximal/non-proximal gradient of *este/non-este*.

2. Materials and methods

2.1. PARTICIPANTS

Similar to Rocca and Wallentin (2020), we conducted a DCT experiment using Qualtrics (http://qualtrics.com) with participants recruited through the website Prolific (http://prolific.ac; see Appendix B for an example of the experiment). 1639 native Spanish speakers participated (gender: 871 female, 746 male, 22 other; age information was collected in bins: 1013 were 18–30 years, 401 were 30–40 years, 161 were 40–50 years, 56 were 50–60 years, and 8 were 60+ years). The study was approved by the Institutional Review Board at Aarhus University.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

2.2. PROCEDURE

The study took on average 719 (std: 479) seconds to complete, and participants were rewarded with 0.9 GBP for participation. Each participant was presented with 60 Spanish words, sampled and translated from a database of 535 English words rated on 65 different semantic dimensions comprising sensory, motor, spatial, temporal, affective, social, and cognitive experiences (Binder et al., 2016). 480 of the 535 words were divided into 8 subsets, and participants were presented with one such subset of words in a pseudo-randomized manner. Translation of the full list of words from English to Spanish is presented in Appendix C. To allow for comparison with English, we used the same words (translated in Spanish). We provided both masculine and feminine forms for both the noun and the demonstrative options (i.e., noun: *professor/a* 'teacher'; demonstrative options: *este/a* 'this'; *ese/a* 'this-that'; *aquel/la* 'that'), or only for the demonstratives when the noun had the same form for both feminine and masculine form (i.e., common gender noun: *paciente* 'patient'; demonstrative options: *este/a* 'this'; *ese/a* 'this-that'; *aquel/la* 'that'). In total, we had 57 occurrences of masculine/feminine words, 26 occurrences of common gender nouns; 188 occurrences of exclusively feminine nouns, and 209 of exclusively masculine nouns. Grammatical features of the nouns were in accordance with the linguistic regulation of the Real Spanish Academy (Real Academia Española: https://www.rae.es). Similar to previous experiments (Rocca et al., 2019a; Rocca & Wallentin, 2020), participants were asked to couple each word with a proximal, medial, or distal spatial demonstrative, such as *este/a*, *ese/a*, or *aquel/la* without further context. Demonstrative gender was congruent with the individual words. Participants were instructed to simply follow their intuition and choose the combination of demonstrative and word they thought fitted best. Subsequently, participants were subjected to a short personality test, and two tests of mood. Analysis of these data falls within the scope of a separate study (ongoing) and will not be reported here.

2.3. MATERIALS

The 65 semantic dimensions that words are rated along in the Binder dataset are: *Vision, Bright, Dark, Colour, Pattern, Large, Small, Motion, Biomotion, Fast, Slow, Shape, Complexity, Face, Body, Touch, Temperature, Texture, Weight, Pain, Audition, Loud, Low, High, Sound, Music, Speech, Taste, Smell, Head, UpperLimb, LowerLimb, Practice, Landmark, Path, Scene, Near, Toward, Away, Number, Time, Duration, Long, Short, Caused, Consequential, Social, Human, Communication, Self, Cognition, Benefit, Harm, Pleasant, Unpleasant, Happy, Sad, Angry, Disgusted, Fearful, Surprised, Drive, Needs, Attention,* and *Arousal*. The database is publicly available here <http://www.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

neuro.mcw.edu/representations/index.html>, and the rationale for the choice of these exact features is that they represent "experiential phenomena for which there are likely to be corresponding distinguishable neural processors" (Binder et al., 2016, pp. 133–134). The notion that these features should have clearly defined neural underpinnings suggests that they are somehow important and representative for human cognition (see Binder et al., 2016, for further details).

Similar to Rocca and Wallentin (2020), we also included the Lancaster Sensorimotor Norms (available here <https://osf.io/7emr6/>) to our feature set to improve the semantic profile on manipulability. This dataset provides ratings along 11 sensorimotor features for a large body of words (Lynott, Connell, Brysbaert, Brand, & Carney, 2020). The 11 dimensions are the following (the affix *Lan* is appended to differentiate from features from the Binder dataset): *Auditory_Lan, Gustatory_Lan, Haptic_Lan, Interoceptive_Lan, Olfactory_Lan, Visual_Lan, Torso_Lan, Foot_leg_Lan, Hand_arm_Lan, Head_Lan, and Mouth_Lan* (see Rocca and Wallentin, 2020, for illustrations of the features ordered according to semantic factors, obtained by factor analysis).

The overlap between the stimulus words and the two databases included 472 out of the original 480 words. All feature ratings were standardized to make them comparable. Two Binder features contained missing ratings for particular words. These were imputed using the mean of all other words along that feature.

2.4. STATISTICAL ANALYSES

Analyses were conducted in RStudio version 1.1.383 (RStudio Team, 2016) running R version 4.0.2 (R Core Team, 2020). Additional packages used were: tidyverse 1.3.0 (Wickham et al., 2019), xlsx 0.6.3 (Dragulescu & Arendt, 2020), psych 2.0.7 (Revelle, 2017), ggsci 2.9 (Xiao & Li, 2018), cowplot 1.0.0 (Wilke, 2020), reshape2 1.4.4 (Wickham, 2020), scales 1.1.1 (Wickham, Seidel, & RStudio, 2020), and MBA 0.0.9 (Finley, Banerjee, Hjelle, & Bivand, 2017).

The analysis involved the following steps made in order to replicate the analyses conducted by Rocca and Wallentin (2020) while also adapting the procedure to a three-demonstrative language:

1. Exploratory factor analysis on independent variables (semantic features) to lower dimensionality and create independent semantic dimensions (identical to Rocca & Wallentin, 2020).
2. Aggregating demonstrative choices, i.e., determining proportion of the choice of *este/a*, *ese/a*, and *aquel/la* for each word. Conducting principal component analysis on the proportions. This turns the three options into two dimensions (i.e., two principal components (PCs)), which correspond to the effective degrees of freedom in a three-choice experiment. This

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

adapts the analyses conducted by Rocca and Wallentin (2020) to the Spanish case.

3. Conducting two regression analyses, one for each PC, using semantic factors to predict the PC scores for each word.

Each of the steps is described in detail below. Data and analysis scripts are available at the Open Science Framework (https://osf.io/hjgfb/).

2.5. EXPLORATORY FACTOR ANALYSIS OF SEMANTIC FEATURES

We used the same dimensionality reduction approach to the semantic space (i.e., the 76 semantic features) as that used by Rocca and Wallentin (2020). This was aimed at lowering the number of correlated regressors to be used in statistical analyses while preserving structural factors of the semantic space. To determine the number of latent factors, we used Horn's parallel method (Horn, 1965), implemented in the psych package (Revelle, 2019) in R. This method compares the scree plot from the observed data with one made from random samples (randomized across rows) of the original data, and subtracts out the components that explain less variance than a comparable factor based on non-informative data. The estimated number of non-random factors in the semantic features using this procedure was 12. Factor analysis was conducted using Ordinary Least Squares (OLS) to find the minimum residual (minres) solution. Orthogonal rotation (varimax) was applied to maintain orthogonality among factors (important for use in regression analyses). The cumulative proportion of variance of the semantic features explained by the 12 factors was 0.75.

The same factor labels as those used by Rocca and Wallentin (2020) were applied (see Figures 1–3). The 12 factors and the proportion of the variance they explained in the original semantic space were: *Vision* (0.14), *Valence* (0.11), *Loudness* (0.09), *Human* (0.06), *Taste/Smell* (0.06), *Motion* (0.06), *Manipulability* (0.06), *Scene* (0.05), *Time* (0.03), *Torso/Legs* (0.03), *Arousal* (0.03), *Self* (0.03). It is important to note that these factors and the relative variance they explain do not necessarily reflect the general distribution in language or semantics, but only in the underlying sample of words and features present in the combined Binder and Lancaster databases. The structure and ordering of the factors are therefore also partly specific to those words. The factors may nevertheless serve to investigate whether semantics plays a role in demonstrative choices, and given that the same factors and words were previously used in an analysis of English demonstratives, a comparison may show similarities and differences between the languages. The 12 factors were therefore used as predictors in two regression analyses, where they were used to predict demonstrative choices.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

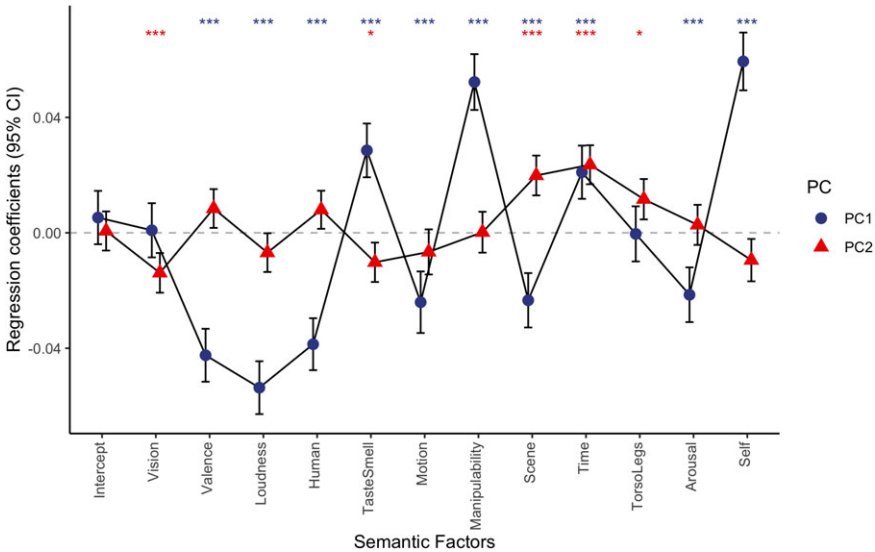


Fig. 1. Regression coefficients (with 95% confidence intervals) for the two regression analyses using semantic factors to predict percentage demonstrative choices for words. Factor analysis on a combination of 76 Binder and Lancaster semantic features resulted in 12 factors, depicted on the *x*-axis. Factors are labelled according to Rocca and Wallentin (2020). Demonstrative choice of *este/a* 'this', *ese/a* 'this/that', and *aquel/la* 'that' were condensed to two principal components, PC1 (blue circle) corresponds to the choice of *este/a* 'this' vs. the two non-proximal demonstratives; PC2 (red triangles) corresponds to *ese/a* 'this/that' vs. *aquel/la* 'that'. A significant positive coefficient for PC1 means that words with a high score on a sematic factor are more likely to elicit the proximal *este/a*, whereas a negative coefficient means that a factor tends to elicit one of the two non-proximal demonstratives more often indiscriminately. A positive coefficient for PC2 means that a word with a high score on a semantic factor is more likely to elicit the distal *aquel/la*, whereas a negative coefficient on PC2 means that a factor is more likely to elicit the medial *ese/a*. * $p < .05$, ** $p < .01$, *** $p < .001$, Bonferroni corrected.

2.6. AGGREGATE COMPONENTS OF DEMONSTRATIVE CHOICES AND PRINCIPAL COMPONENT ANALYSES

Individual choices of demonstrative for words were aggregated across participants to yield a proportion of *este/a* 'this', *ese/a* 'this/that', and *aquel/la* 'that' chosen for each word (see Table 1 for examples and overall proportions). Given that the proportions always sum to 100%, the distribution across the three demonstratives has only two degrees of freedom. Using principal component analysis (PCA), we therefore projected the three different demonstrative options onto the two dimensions (PC1 and PC2) on which choices were distributed. The loadings for these two components across the three demonstratives were for PC1: [0.81, –0.34, –0.47], i.e., roughly corresponding to the proximal demonstrative *este/a* relative to the two other demonstratives combined. For PC2 it was: [0.08, –0.74, 0.66], i.e., roughly corresponding to the

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

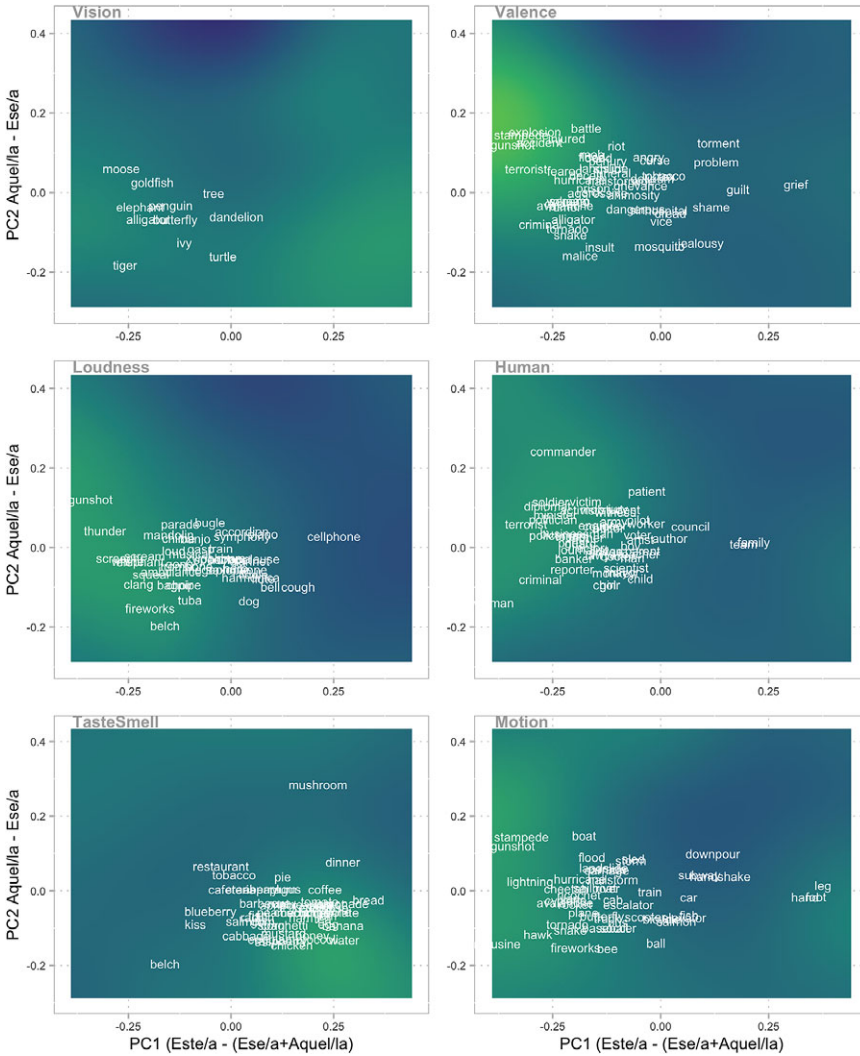


Fig. 2. Each plot shows how demonstrative choices are distributed for semantic factors 1–6 (based on 76 Binder and Lancaster features): Vision, Valence, Loudness, Human, Taste/Smell, Motion. The *x*-axis represents principal component 1 for demonstrative choices: a positive number means that participants more often use *este/a* 'this' than any of the other two demonstratives for words that score high on a particular semantic factor. The *y*-axis displays principal component 2: the difference between *aquel/la* 'that' and *ese/a* 'this/that'. A positive number means that a*quel/la* 'that' is chosen more often for words related to this factor. We used multilevel B-splines (Lee, Wolberg, & Shin, 1997) to interpolate the distribution of each semantic factor in 2D demonstrative space (surface plot). Yellow colours indicate a higher frequency of use for the semantic factor. The plots also include high-scoring words on each semantic factor (cut-off 1.66 standard deviations above the mean) for illustration (Spanish version in Appendix A).

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

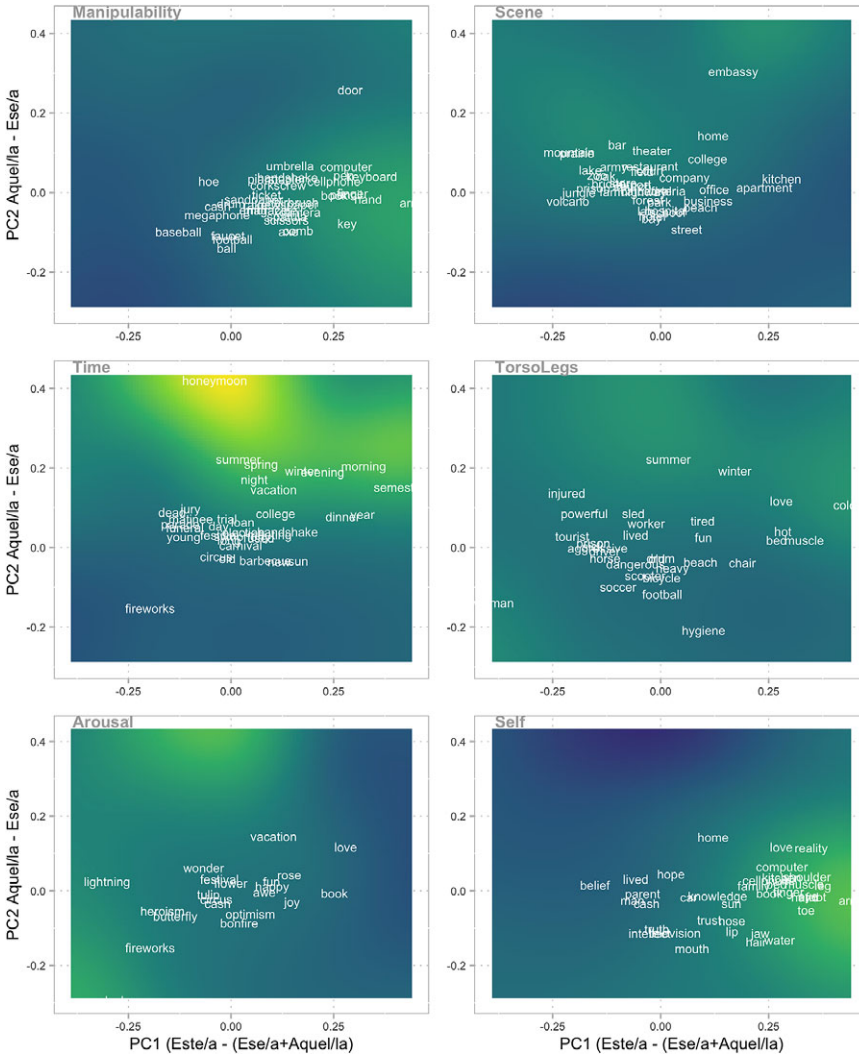


Fig. 3. Each plot shows how demonstrative choices are distributed for semantic factors 7–12 (based on 76 Binder and Lancaster features): Manipulability, Scene, Time, Torso/Legs, Arousal, Self. Yellow colours indicate a higher frequency of use for the semantic factor. The *x*-axis represents principal component 1 for the demonstrative choices: a positive number means that participants more often use *este/a* 'this' than any of the other two demonstratives for words that score high on a particular semantic factor. The *y*-axis represents principal component 2: the difference between *aquel/la* 'that' and *ese/a* 'this/that'. A positive number means that *aquel/la* 'that' is chosen more often for words related to this factor. We used multilevel B-splines (Lee et al., 1997) to interpolate the distribution of each semantic factor in 2D demonstrative space (surface plot). The plots also include high-scoring words on each semantic factor (cut-off 1.66 standard deviations above the mean) for illustration (Spanish version in Appendix A).

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

distal demonstrative *aquel/a* relative to the medial demonstrative *ese/a*, disregarding the proximal demonstrative *este/a*.

2.7. REGRESSION ANALYSES

We conducted a linear regression analysis for each of the two principal components (PC1 and PC2) with PC scores for each word as dependent variable and estimated factor scores for the 12 factors as independent variables. This allowed us to determine which (if any) semantic factors could be used to predict the distribution of demonstrative choices across the two PC dimensions. Results were corrected for multiple comparisons using a Bonferroni correction.

3. Results

The overall proportion of the three demonstratives can be seen from Table 1, along with distributions for example words.

3.1. PRINCIPAL COMPONENT 1

The first linear regression model with semantic factors as independent variables and PC1 as dependent variable was highly significant (Adjusted R-squared: 0.584), indicating that the semantic factors explained variability

TABLE 1. *Example words from the study*

Example words		Proportion demonstrative			Principal components	
English	*Spanish*	*Este/a* (this)	*Ese/a* (this/ that)	*Aquel/la* (that)	*PC1*	*PC2*
ambulance	ambulancia	0.25	0.45	0.30	–0.145	–0.05
banana	plátano	0.58	0.34	0.08	0.27	–0.08
diplomat	diplomático	0.16	0.37	0.47	–0.26	–0.12
duck	pato/a	0.24	0.49	0.27	–0.15	–0.10
family	familia	0.55	0.27	0.18	0.21	0.02
home	casa	0.48	0.21	0.31	0.12	0.15
night	noche	0.43	0.21	0.36	0.06	0.18
toe	dedo del pie	0.64	0.28	0.08	0.34	–0.04
…	…	…	…	…	…	…
Mean (480 words)		***0.37***	***0.37***	***0.26***		

NOTE. Example words from the study. Participants had to choose one of the three spatial demonstratives that they thought was best for that word without context. A proportion of choices were condensed into two principal components which were used as dependent variables in regression analyses. The mean value reported in the table refers to the total number of words in the experiment.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

in the choice between the proximal demonstrative *este/a* and the two other demonstratives (i.e., medial *ese/a* and distal *aquel/la*).

Out of the 12 semantic factors, 10 significantly contributed to the model ($p <$ .05, Bonferroni corrected, see Figure 1): *Valence* ($t(459) = -9.1$, $p < .0001$), *Loudness* ($t(459) = -11.5$, $p < .0001$), *Human* ($t(459) = -8.4$, $p < .0001$), *Taste/Smell* ($t(459) = 6.0$, $p < .001$), *Motion* ($t(459) = -4.4$, $p < .0002$), *Manipulability* ($t(459) = 10.6$, $p < .0001$), *Scene* ($t(459) = -4.9$, $p < .0001$), *Time* ($t(459) = 4.5$, $p < .0001$), *Arousal* ($t(459) = -4.4$, $p < .0002$), and *Self* ($t(459) = 11.6$, $p <$ .0001). The factors *Vision* and *Torso/Legs* were non-significant ($p > .05$). Positive *t*-values indicate that the factor contributes positively to PC1 (i.e., elicits the proximal demonstrative *este/a* more often), whereas negative coefficients and *t*-values indicate a negative contribution to PC1 (i.e., elicits the medial *ese/a* or distal demonstrative *aquel/la* more often).

3.2. PRINCIPAL COMPONENT 2

The second linear regression model with semantic factors as independent variables and PC2 as dependent variable was also highly significant (Adjusted R-squared: 0.22), indicating that the semantic factors explained variability in the choice between the distal demonstrative *aquel/la* and the medial demonstrative ese/*a*. Out of the 12 semantic factors, 5 significantly contributed to the model ($p < .05$, Bonferroni corrected, see Figure 1): *Vision* ($t(459) = -3.9$, $p < .0001$), *Taste/Smell* ($t(459) = -2.9$, $p < .05$), *Scene* ($t(459) = 5.6$, $p < .0001$), *Time* ($t(459) = 6.8$, $p < .0001$), *Torso/Legs* ($t(459) = 3.2$, $p < .05$). The remaining factors were non-significant ($p > .05$) when correcting for multiple comparisons. Positive *t*-values indicate that the factor contributes positively to PC2 (i.e., elicits the distal demonstrative *aquel/la* more often), whereas negative coefficients and *t*-values indicate a negative contribution to PC2 (i.e., elicits the medial demonstrative *ese/a* more often).

4. Discussion

In the present study, we adapted the DCT to Spanish to assess the influence of semantic features of the referent on speakers' choices for demonstrative forms in a three-term demonstrative system (*este/ese/aquel* 'this/this-that/that' in English), when no contextual cue is provided. Our aim was twofold. First, we aimed at ascertaining whether results obtained with a two-terms demonstrative system replicate in a three-term demonstrative system, that is, whether the same semantic dimensions influence speakers' preferences for proximal vs. non-proximal demonstrative forms when a language does not present a dyadic demonstrative system. Second, we aimed to assess whether the contrast between non-proximal (i.e., medial and distal) forms can also be used for

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

SEMANTIC distinctions, that is, to assess whether particular semantic features of the referent explain speakers' choices of medial vs. distal forms, or whether the distinction between medial and distal forms is only relevant in pragmatic contexts.

Concerning the first question, our results replicate the findings from a previous DCT study with English, showing that a range of semantic factors consistently influences demonstrative choice when no context is provided (Rocca & Wallentin, 2020). The same 10 semantic factors (including Manipulability, Valence, and Self) that predicted the distribution of proximal and distal demonstrative choices in English ALSO predicted the choice of proximal (*este/a*) vs. non-proximal (*ese/a – aquel/la*) demonstratives in Spanish (PC1; see Figure 1).

Concerning the second question, we outlined two potential hypotheses. Hypothesis A stated that, in the absence of any context, the three-system demonstrative system might collapse into a two-term (proximal/non-proximal) demonstrative system organized along a self/non-self dichotomy, where medial forms do not have functional/semantic specificity. Hypothesis B, on the other hand, stated that medial and distal forms retain some semantic specificity with respect to each other, and consequently that some semantic factor would explain speakers' choice for one form rather than the other.

Our results speak against hypothesis A. If the three-term system was reduced to a two-term system, one could either expect the use of a medial or distal form to be drastically reduced (which was not the case, as it can be seen from Table 1) or that speakers would use the medial and the distal term more or less interchangeably, and therefore randomly. In this scenario, semantic factors should NOT be consistently related to participants' choices for medial (*ese/a*) versus distal (*aquel/la*) forms. However, we DO detect a relationship between particular semantic dimensions of the referent and the speakers' tendency to choose medial versus distal forms. In fact, PC2, which modelled the difference in use between the medial *ese/a* and the distal *aquel/la*, was significantly correlated with five different semantic factors (see Figure 1). This lends support to hypothesis B, which stated that, in the absence of spatial information, the dichotomy between the two non-proximal forms can be used to encode specific semantic distinctions.

An especially prominent example of such a semantic factor is Time (see Figures 1 and 3). As can be seen from Figure 3 (row 2, column 1), words related to Time caused a significant increase in the frequency of the distal form *aquel/la* compared to the medial form *ese/a*. Participants tended to choose *aquella luna de miel* instead of *esa luna de miel* 'that honeymoon' and *aquel verano* instead of *ese verano* 'that summer'. Similar results have been found for other languages, such as Estonian, a two-term system with the forms *see* 'this-that' and *too* 'that', where the distal term *too* 'that' has been found to be especially used for fixed

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

expressions related to the same semantic dimension of Time (Pajusalu, 2006). Conversely, a high score on the factor Vision increased participants' likelihood to use the medial demonstrative *ese/a* relative to the distal demonstrative *aquel/la*. As can be seen from Figure 2, high scorers on this factor include animate beings, such as animals and plants. Participants tended to use *ese tigre* instead of *aquel tigre* 'that tiger', and *esa tortuga* was more frequent than *aquella tortuga* 'that turtle'. Animacy has previously been found to influence choice of demonstratives (Rocca et al., 2019a), especially in interaction with harmfulness. Isolating the individual contribution of Animacy (for which no explicit factor exists in this analysis) vs. Vision requires more fine-grained analyses. Overall, our results thus suggest that, while in Spanish the proximal/non-proximal distinction is explained by the same semantic factors detected for the proximal/distal contrast in English (with the same 10 semantic factors contributing to demonstrative choice), in Spanish the medial factor also retains semantic specificity.

Our observations that *ese/a* 'this-that' and *aquel/la* 'that' may be used to code for semantic distinctions are in line with observations from Argentinian Spanish, where patterns of usage of medial vs. distal terms are also consistently driven by non-spatial dimensions (here, *aquel/la* is used only in formal contexts). Another parallel is Majorcan Catalan, where the medial term *aqueix/a* has dropped out of the system when conveying physical distance, but can be used to mark social distance (Pérez-Saldanya, 2015; Todisco et al., 2021). A general description of the relationship between expressiveness and complexity (i.e., number of forms) of demonstrative systems goes beyond the scope of the present studies, but it is an extremely interesting research avenue. On the one hand, the examples of Spanish and Argentinian Spanish may indeed suggest that maintaining a three-term (more complex) system may allow more refined referencing in relevant situations, also outside physical space. On the other hand, in languages with 'simpler' two-term demonstrative referencing systems, modulations of the dyadic system (e.g., by combining demonstrative particles and adverbs) may be used to carve more fine-grained distinctions and therefore retain high expressivity. Italian, for instance, presents a two-term demonstrative system (*questo/quello*), but speakers often use these combinations of a demonstrative and a locative adverb for more fine-grained referencing (e.g., *questo qui* 'this here'; Jungbluth & Da Milano, 2015). Studies on the relationships and trade-offs between lexical complexity of demonstrative systems, affordances for pragmatic modulations, and expressivity would, in our opinion, be a highly valuable contribution to the literature.

To conclude, showing that semantic patterns psychologically or quantitatively replicate not only in languages with two-term demonstrative systems (i.e., Italian, Danish in Rocca et al., 2019a; English in Rocca & Wallentin, 2020), but also in languages with more than two demonstrative terms provides

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

a strong argument in favour of these mechanisms capturing some fundamental aspect of the relation between human cognition, semantics, and linguistic forms. Overall, these results are in line with the idea suggested by Rocca and Wallentin (2020), that the frame of reference for demonstrative use – at least in absence of an explicit interlocutor – instantiates the speaker as deictic centre, and projects away from this centre in both physical and psychological dimensions. When non-spatial semantic entities are referred to, the coordinate system is structured along psychological (e.g., familiarity, affect), semantic, and imaginative dimensions, rather than spatial ones (Coventry et al., 2014; Stukenbrock, 2014; Rocca & Wallentin, 2020).

A number of points remain open to investigation. First, our results lend support to the hypothesis that semantics systematically influences the choice of specific demonstrative forms when no context is provided. So far, however, effects of semantic dimensions on demonstrative choice have been described for a few languages, and some that are arguably not so dissimilar from a cultural and typological standpoint. A more extensive overview of the extent to which the detected dynamics generalize across languages is warranted. Conducting the DCT on a wider range of languages (using comparable words and feature sets), ideally covering a broader typological and cultural spectrum, would pave the way towards a deeper understanding of the cognitive mechanisms underlying the complex functioning of deictic reference and its relationship with cultural differences, given that the semantic dimensions will be the same for each language under analysis, independently of the context of utterance.

Investigating how this semantic space interacts with physical and conversational space is also a crucial avenue for future research, together with a more in-depth analysis of collocation patterns found in corpora of spoken and written language. This may shed light on whether the same semantic distinctions observed through the DCT can be observed in 'naturalistic' contexts. The analysis of corpora may replicate the effects found in our studies, but also enrich the picture with knowledge of usage patterns which are more conventional (e.g., common colocations) than purely driven by semantics.

Finally, our understanding of the topology of demonstrative systems with more than two terms and their relation to a dyadic system could be significantly improved by expanding the DCT to languages, such as Catalan, which present a reduction from three to two terms.

5. Conclusions

In the present study, we used the Demonstrative Choice Task (DCT) to investigate the relationship between demonstrative use and semantics in Spanish, a language with three demonstrative terms. We have provided evidence that demonstrative choice is influenced by a number of semantic dimensions

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

(spatial, bodily, and emotional features). The contrast between proximal and distal forms is driven by the same dimensions which have been previously found to influence demonstrative choice in English. Semantic effects are thus replicable across languages, which hints at some degree of invariance in cognitive mechanisms underlying demonstrative reference. Additional semantic factors consistently influenced speakers' preferences for medial and distal forms, which suggests that the medial/distal distinction can be recruited to express finer-grained semantic contrasts.

REFERENCES

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M. & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology* **33**(3/4), 130–174.

Bühler, K. (1934/2011). *Theory of language: the representational function of language* [*Sprachtheorie*]. Amsterdam: John Benjamins.

Bocale, P. & Cologna, D. B. (2020). The expression of epistemic uncertainty through the use of distal demonstratives in Russian and Chinese: a cognitive analysis of corpus data. *International Journal of Linguistics* **12**(2), 57–82.

Caldano, M. & Coventry, K. R. (2019). Spatial demonstratives and perceptual space: To reach or not to reach? *Cognition* **191**, e2019.06.001.

Cooperrider, K. (2016). The co-organization of demonstratives and pointing gestures. *Discourse Processes* **53**, 632–656.

Coventry, K. R., Griffiths, D. & Hamilton, C. J. (2014). Spatial demonstratives and perceptual space: describing and remembering object location. *Cognitive Psychology* **69**, 46–70.

Coventry, K. R., Valdés, B., Castillo, A. & Guijarro-Fuentes, P. (2008). Language within your reach: near–far perceptual space and spatial demonstratives. *Cognition* **108**(3), 889–895.

Di Pellegrino, G. & Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia* **66**, 126–133.

Diessel, H. (1999). *Demonstratives: form, function and grammaticalization.* Amsterdam: John Benjamins.

Diessel, H. (2005). Distance contrasts in demonstratives. In M. Haspelmath, M. S. Dryer, D. Gil, D. & B. Comrie (eds), *World atlas of language structures* (pp. 170–173). Oxford: Oxford University Press.

Diessel, H. (2013). Where does language come from? Some reflections on the role of deictic gesture and demonstratives in the evolution of language. *Language and Cognition* **5**(2/3), 239–249.

Diessel, H. (2014). Demonstratives, frames of reference, and semantic universals of space. *Language and Linguistics Compass* **8**(3), 116–132.

Diessel, H. & Coventry, K. R. (2020). Demonstratives in spatial language and social interaction: an interdisciplinary review. *Frontiers of Psychology* **11**, e555265.

Dragulescu, A. & Arendt, C. (2020). xlsx. Online <https://CRAN.R-project.org/package=xlsx>.

Dufter, A. (2015). Semantics. In K. Jungbluth & F. Da Milano (eds), *Manual of deixis in Romance languages* (pp. 359–380). Berlin/Boston: Walter de Gruyter.

Finley, A., Banerjee, S., Hjelle, Ø. & Bivand, R. (2017). MBA: Multilevel B-Spline Approximation. Online <https://CRAN.R-project.org/package=MBA>.

Fricke, E. (2014). Deixis, gesture, and embodiment from a linguistic point of view. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill & J. Bressem (eds), *Body, language, communication* (pp. 1803–1823). Berlin: De Gruyter Mouton.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Gudde, H. B., Griffiths, D. & Coventry, K. R. (2018). The (spatial) memory game: testing the relationship between spatial language, object knowledge, and spatial cognition. *Journal of Visualized Experiments* **132**, e56495.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika* **30**(2), 179–185.

Hottenroth, P. (1982). *Locative in Spanish, in here and there: cross-linguistics studies on deixis and demonstratives*. Amsterdam: John Benjamin.

Hunley, S. B. & Lourenco, S. F. (2018). What is peripersonal space? An examination of unresolved empirical issues and emerging findings. *Wiley Interdisciplinary Reviews, Cognitive Science* **9**(6), e1472.

Iverson, J. M., Tencer, H. L., Lany, J. & Goldin-Meadow, S. (2000). The relation between gesture and speech in congenitally blind and sighted language-learners. *Journal of Nonverbal Behavior* **24**(2), 105–130.

Jungbluth, K. (2003). Deictics in the conversational dyad: findings in Spanish and some cross-linguistic outlines. In F. Lenz (ed.), *Deictic conceptualization of space, time and person* (pp. 13–40). Amsterdam: John Benjamins.

Jungbluth, K. & Da Milano, F. (2015). *Manual of deixis in Romance languages*. Berlin/Boston: Walter de Gruyter.

Kemmerer, D. (2006). The semantics of space: integrating linguistic typology and cognitive neuroscience. *Neuropsychologia* **44**(9), 1607–1621.

Kita, S. (2003). *Pointing: where language culture, and cognition meet*. Mahwah, NJ: Lawrence Erlbaum Associates.

Küntay, A. C. & Özyürek, A. (2002). Joint attention and the development of the use of demonstratives in Turkish. *Proceedings of the 26th Annual Boston University Conference on Language Development 26* (pp. 336–347). Somerville, MA: Cascadilla Press.

Lee, S., Wolberg, G. & Shin, S. Y. (1997). Scattered data interpolation with multilevel B-splines. *IEEE Transactions on Visualization and Computer Graphics* **3**, 228–244.

Leech, G., Rayson, P. & Wilson, A. (2014). *Word frequencies in written and spoken English: based on the British National Corpus*. London: Routledge.

Levinson, S. (2018). Introduction: patterns in diversity. In S. Levinson, S. Cutfield, M. Dunn, N. Enfield & S. Meira (eds), *Demonstratives in cross-linguistic perspective (Language Culture and Cognition)* (pp. 1–42). Cambridge: Cambridge University Press.

Levinson, S., Cutfield, S., Dunn, M., Enfield, N. & Meira, S. (eds) (2018). *Demonstratives in cross-linguistic perspective (Language Culture and Cognition)*. Cambridge: Cambridge University Press.

Lin L. (2013). *A semantic study of German and Chinese demonstratives*. Unpublished doctoral dissertation, University of Los Angeles, California. Online <https://escholarship.org/uc/item/9v40c1w2>.

Lynott, D., Connell, L., Brysbaert, M., Brand, J. & Carney, J. (2020). The Lancaster Sensorimotor Norms: multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods* **52**, 1271–1291.

Pajusalu, R. (2006). Death of a demonstrative: person and time. The case of Estonian too. *Linguistica Uralica* **42**(4), 241–253.

Peeters, D., Hagoort, P. & Özyürek A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition* **136**, 64–84.

Peeters, D., Krahmer, E. & Maes, A. A. (2021). A conceptual framework for the study of demonstrative reference. *Psychonomic Bulletin & Review* **28**, 409–433.

Pérez-Saldanya, M. (2015). Paradigms as triggers of semantic change: demonstrative adverbs in Catalan and Spanish. *Catalan Journal of Linguistics* **14**, 113–135.

R Core Team (2020). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Online <https://www.R-project.org/>.

Revelle, W. (2017). psych: procedures for personality and psychological research. Northwestern University, Evanston, IL. *Online* <https://CRAN.R-project.org/package=psych>.

Revelle, W. (2019). psych: procedures for psychological, psychometric, and personality research. Northwestern University, Evanston, IL.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Rocca, R., Tylén, K. & Wallentin, M. (2019a). This shoe, that tiger: semantic properties reflecting manual affordances of the referent modulate demonstrative use. *PloS One* **14**(1), e0210333.
Rocca, R. & Wallentin, M. (2020). Demonstrative reference and semantic space: a large-scale Demonstrative Choice Task study. *Frontiers in Psychology*, e2020.00629.
Rocca, R., Wallentin, M., Vesper, C. & Tylén, K. (2018). This and that back in context: grounding demonstrative reference in manual and social affordances. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*. Madison, WI. Online <https://cogsci.mindmodeling.org/2018/papers/0192/0192.pdf>.
Rocca, R., Wallentin, M., Vesper, C. & Tylén, K. (2019b). This is for you: social modulations of proximal vs. distal space in collaborative interaction. *Scientific Reports* **9**(1), 1–14.
RStudio Team. (2016). *RStudio: integrated development for R*. RStudio, Boston: MA. Online <http://www.rstudio.com/>.
Strauss, S. (2002). This, that, and it in spoken American English: a demonstrative system of gradient focus. *Language Sciences* **24**(2), 131–152.
Stukenbrock, A. (2014). Pointing to an 'empty' space: *deixis am phantasma* in face-to-face interaction. *Journal of Pragmatics* **74**, 70–93.
Stukenbrock, A. (2015). Intercorporeal phantasms: kinesthetic alignment with imagined bodies in self-defense trainings. *International and Linguistic Structures* **58**, 1–23.
Talmy, L. (1983). How language structures space. In H. Pick (ed.), *Spatial orientation: theory, research, and application* (pp. 225–282). Boston, MA: Springer.
Talmy, L. (1988). The relation of grammar to cognition. In B. Rudska-Ostyn (ed.), *Topics in cognitive linguistics* (pp. 165–206). Amsterdam: John Benjamins.
Talmy, L. (2018). *The targeting system of language*. Cambridge, MA/London: MIT Press.
Todisco, E., Guijarro-Fuentes, P., Collier, J. & Coventry, K. R. (2020). The temporal dynamics of deictic communication. *First Language* **41**(2), 154–178.
Todisco, E., Guijarro-Fuentes, P. & Coventry, K. R. (2021). Analogical levelling in the Majorcan Catalan demonstrative system. *Probus International Journal of Romance Linguistics*.
Wickham, H. (2020). reshape2. Online <https://cran.r-project.org/web/packages/reshape2/index.html>.
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., … & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software* **4**(43), e01686.
Wickham, H., Seidel, D. & RStudio (2020). scales. Online <https://cran.r-project.org/web/packages/scales/index.html>.
Wilke, C.O. (2020). cowplot. Online <https://cran.r-project.org/web/packages/cowplot/index.html>.
Wilkins, D. P. (1999). Demonstrative questionnaire: "THIS" and "THAT" in comparative perspective. In D. P. Wilkins (ed.), *Manual for the 1999 field season* (pp. 1–24). Nijmegen: MPI for Psycholinguistics.
Xiao, N. & Li, M. (2018). ggsci-package. Online <https://cran.r-project.org/web/packages/ggsci/vignettes/ggsci.html>.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Appendix A

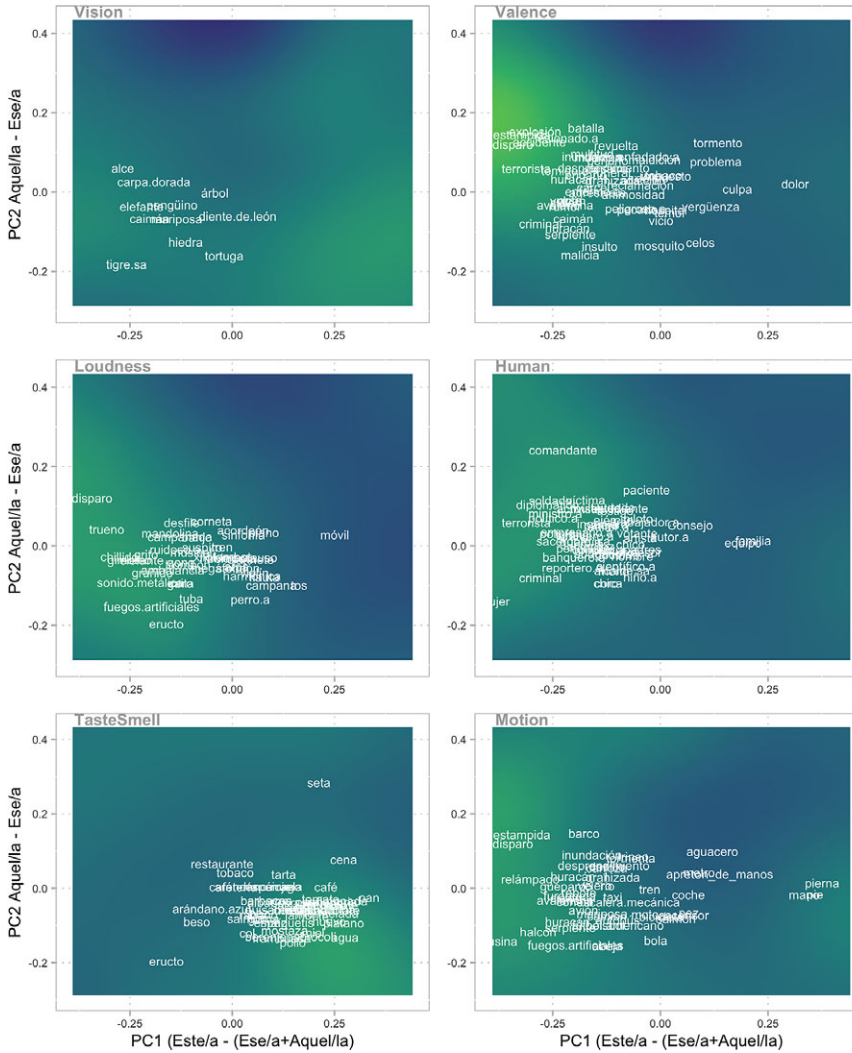


Fig. 4. Demonstrative choice distribution (factor 1-6_Spanish version).

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

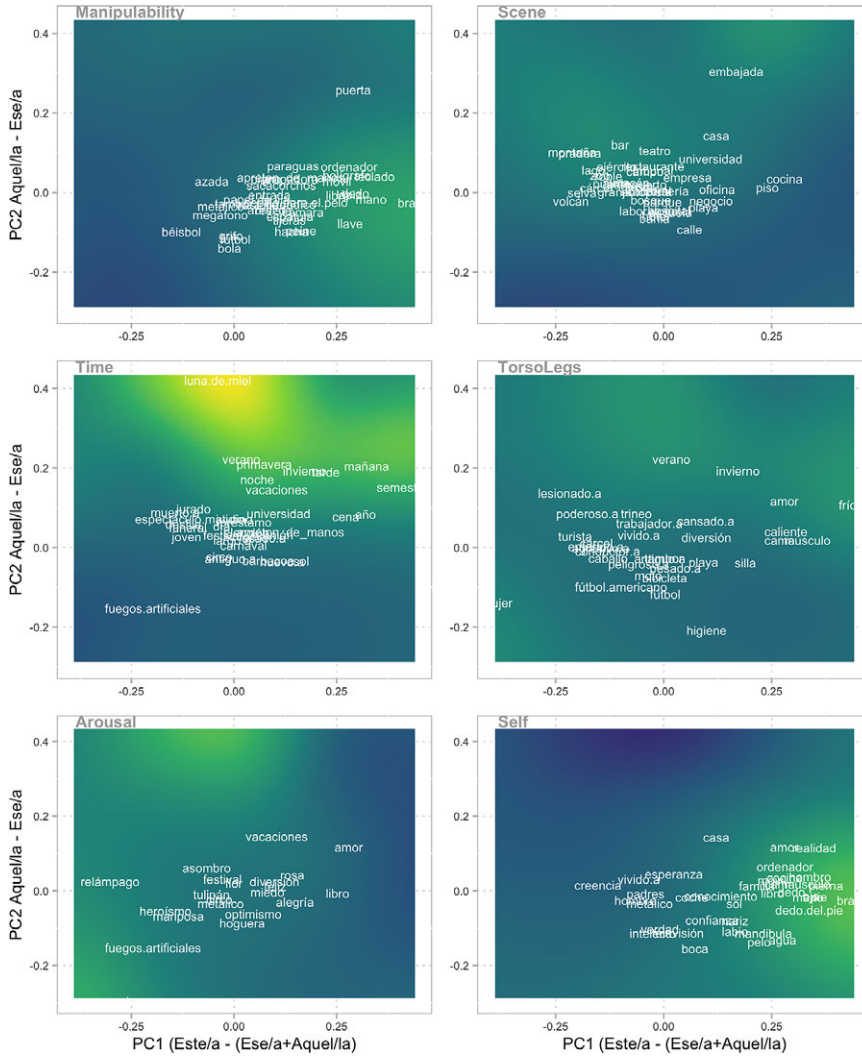


Fig. 5. Demonstrative choice distribution (factor 7-12_Spanish version).

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Appendix B

The figures represent the experimental setting in both laptop and mobile versions.



Fig. 6. Laptop version.



Fig. 7. Mobile version.

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Appendix C

Here we report the words used in the DCT both in English and Spanish.

Word_English	Word_Spanish	Word_English	Word_Spanish
helped	ayudado/a	restaurant	restaurante
advantage	ventaja	ricochet	rebote
applause	aplauso	semester	semestre
awe	miedo	Sin	pecado
banjo	banjo	snub	desaire
banker	banquero/a	spring	primavera
belief	creencia	teacher	profesor/a
bicycle	bicicleta	thunder	trueno
blue	azul	tiger	tigre/sa
bread	pan	tomato	tomate
bridge	puente	truce	tregua
cabinet	gabinete	trumpet	trompeta
cheese	queso	volcano	volcán
circus	circo	white	blanco/a
clever	inteligente	whole	totalidad
clue	pista	xylophone	xilófono
cough	tos	aggressive	agresivo/a
council	Consejo	angry	enfadado/a
cucumber	pepino	barbecue	barbacoa
dark	oscuro/a	bell	campana
day	día	book	libro
driver	conductor/a	business	negocio
farmer	granjero/a	Cab	taxi
faucet	grifo	cabbage	col
feared	temido/a	cafeteria	cafetería
field	campo	carrot	zanahoria
forest	bosque	chair	silla
glass	vaso	church	iglesia
hand	mano	clang	sonido metálico
hoe	azada	clarinet	clarinete
honeymoon	luna de miel	coffee	café
hope	esperanza	dead	muerto/a
hospital	hospital	doctor	doctor/a
hotel	hotel	eggplant	berenjena
ice	hielo	embrace	abrazo
lake	lago	envy	envidia
mushroom	seta	etiquette	etiqueta
old	antiguo/a	excuse	excusa
park	parque	farm	granja
patient	paciente	finger	dedo
piano	piano	fireworks	fuegos artificiales
pig	cerdo/a	football	fútbol
plum	ciruela	gong	gong
reporter	reportero/a	guard	guardia
guilt	culpa	dusty	polvoriento/a
hair	pelo	election	elección
home	casa	escalator	escalera mecánica
jealousy	celos	evening	tarde

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Continued

Word_English	Word_Spanish	Word_English	Word_Spanish
lawyer	abogado/a	expensive	caro/a
magazine	revista	friendly	simpático/a
moose	alce	Fun	diversión
moral	moral	garden	jardín
musical	musical	green	verde
new	nuevo/a	grief	dolor
paradox	paradoja	grievance	reclamación
pen	bolígrafo	happy	feliz
pineapple	piña	harp	arpa
plane	avión	highway	autopista
plea	petición	Hot	caliente
priest	sacerdote/isa	hygiene	higiene
raspberry	frambuesa	Jaw	mandíbula
sailboat	velero	joviality	jovialidad
saxophone	saxofón	judge	juez/a
screech	chillido	limousine	limusina
shame	vergüenza	loud	ruidoso/a
tangerine	mandarina	luck	suerte
turtle	tortuga	malice	malicia
verb	verbo	mayor	alcalde/sa
water	agua	megaphone	megáfono
wealthy	rico/a	mercy	misericordia
winter	invierno	minister	ministro/a
witness	testigo	noun	nombre
woman	mujer	oak	roble
year	año	parade	desfile
young	joven	Pie	tarta
zone	zona	policeman	policía
activist	activista	rake	rastrillo
ambulance	ambulancia	rocket	cohete
avalanche	avalancha	rose	rosa
bonfire	hoguera	salmon	salmón
broccoli	brócoli	saw	sierra
businessman	empresario/a	street	calle
camel	camello/a	theme	tema
camera	cámara	tobacco	tabaco
cherry	cereza	Toe	dedo del pie
choir	coro	tornado	huracán
cranberry	arándano rojo	tourist	turista
criminal	criminal	treaty	tratado
curse	maldición	vacation	vacaciones
vice	vicio	sandpaper	papel de lija
worth	valor	scientist	científico/a
accordion	acordeón	shoulder	hombro
apology	disculpa	soccer	fútbol americano
apricot	albaricoque	speech	habla
bay	bahía	stapler	grapadora
beer	cerveza	student	estudiante
cathedral	catedral	submarine	submarino
cheetah	guepardo	subway	metro
college	universidad	theater	teatro

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Continued

Word_English	Word_Spanish	Word_English	Word_Spanish
corkscrew	sacacorchos	theory	teoría
dandelion	diente de león	torment	tormento
dangerous	peligroso/a	voter	votante
delirium	delirio	wanted	querido/a
dictation	dictado	window	ventana
embassy	embajada	worker	trabajador/a
empty	vacío/a	zoo	zoo
explosion	explosión	damaged	dañado/a
fence	valla	airport	aeropuerto
fiddle	violín	analogy	analogía
fish	pez	Arm	brazo
flower	flor	attribute	atributo
flute	flauta	audience	auditorio
fountain	fuente	bus	autobús
gasp	suspiro	Car	coche
goldfish	carpa dorada	carriage	carroza
hall	sala	chestnut	castaña
honey	miel	chicken	pollo
insult	insulto	computer	ordenador
ire	ira	corn	maíz
jam	mermelada	court	tribunal
lab	laboratorio	desk	escritorio
landslide	desprendimiento	egg	huevo
leg	pierna	Eye	ojo
man	hombre	festival	festival
mandolin	mandolina	folly	locura
matinee	espectáculo matutino	foot	pie
patent	patente	funeral	funeral
pea	guisante	girl	chica
pilot	piloto	gum	goma
plot	guion	hailstorm	granizada
protest	protesta	ham	jamón
reality	realidad	heavy	pesado/a
red	rojo/a	heroism	heroísmo
riot	revuelta	hierarchy	jerarquía
horse	caballo	commander	comandante
hurricane	huracán_1	crib	cuna
infinity	infinito	crow	cuervo
irony	ironía	denial	negación
joke	broma	dog	perro/a
joy	alegría	downpour	aguacero
key	llave	dread	temor
lived	vivido/a	drum	tambor
majority	mayoría	duck	pato/a
mob	multitud	editor	editor/a
mountain	montaña	Elm	olmo
parent	padres	engineer	ingeniero/a
power	poder	family	familia
prairie	pradera	fate	destino
prison	cárcel	hamster	hámster
role	rol	hawk	halcón

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Continued

Word_English	Word_Spanish	Word_English	Word_Spanish
scissors	tijeras	island	isla
shiny	brillante	jungle	selva
small	pequeño/a	jury	jurado
snake	serpiente	ketchup	salsa de tomate
squeal	gruñido	kiss	beso
stampede	estampida	kitchen	cocina
stone	piedra	knowledge	conocimiento
sum	suma	legality	legalidad
summer	verano	long	largo/a
team	equipo	lust	lujuria
television	televisión	meeting	reunión
train	tren	morning	mañana
truck	camión	mosquito	mosquito
used	usado/a	mouth	boca
whine	gimoteo	nose	nariz
yellow	amarillo/a	number	número
alligator	caimán	peaceful	pacífico/a
army	ejército	politician	político/a
bagpipe	gaita	radish	rábano
banana	plátano	river	río
bar	bar	shelves	estanterías
bed	cama	soft	suave
blueberry	arándano azul	spaghetti	espaguetis
boy	chico	spatula	espátula
bribe	soborno	spiritual	espiritual
butterfly	mariposa	testimony	testimonio
carnival	carnaval	tired	cansado/a
cash	metálico	trial	juicio
cold	frío	tulip	tulipán
umbrella	paraguas	scream	grito
wit	humor	sick	enfermo/a
accident	accidente	soldier	soldado
advice	sugerencia	Sun	sol
animosity	animosidad	sympathy	piedad
ant	hormiga	symphony	sinfonía
apartment	piso	table	mesa
artist	artista	Tax	impuesto
author	autor/a	Tea	té
baseball	béisbol	terrorist	terrorista
big	grande	ticket	entrada
bird	pájaro/a	tribute	tributo
black	negro/a	trombone	trombón
boat	barco	trust	confianza
bugle	corneta	tuba	tuba
cellphone	móvil	van	furgoneta
chime	campanada	whale	ballena
chipmunk	ardilla	actor	actor/actriz
chocolate	chocolate	agreement	acuerdo
cloud	nube	asparagus	espárrago
comb	peine	Axe	hacha
deceit	engaño	ball	bola

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press

Continued

Word_English	Word_Spanish	Word_English	Word_Spanish
elephant	elefante	battle	batalla
era	era	beach	playa
fee	tarifa	Bee	abeja
gunshot	disparo	belch	eructo
handshake	apretón de manos	child	niño/a
heredity	herencia	company	empresa
keyboard	teclado	couple	pareja
lemonade	limonada	cyclone	ciclón
lonely	solitario/a	debate	debate
love	amor	dime	moneda de 10 céntimos
medicine	medicina	dinner	cena
motive	motivo	diplomat	diplomático
mouse	ratón	dolphin	delfín
muscle	músculo	door	puerta
mustard	mostaza	elevator	ascensor
newspaper	periódico	fallacy	falacia
night	noche	famous	famoso/a
oration	discurso	feather	pluma
peace	paz	flood	inundación
pumpkin	calabaza	gratitude	gratitud
rumo	rumor	hairbrush	cepillo para el pelo
satire	sátira	harmonica	harmónica
school	escuela	injured	lesionado/a
intellect	intelecto	perjury	perjurio
ivy	hiedra	powerful	poderoso/a
journalist	periodista	problem	problema
law	ley	quantity	cuantidad
lightning	relámpago	rally	rally
liked	apreciado/a	rum	ron
lip	labio	scooter	moto
loan	préstamo	sled	trineo
monkey	mono	store	almacén
mystery	misterio	storm	tormenta
office	oficina	toaster	tostadora
optimism	optimismo	tree	árbol
pan	sartén	truth	verdad
party	fiesta	victim	víctima
pencil	lápiz	woe	aflicción
penguin	pingüino	wonder	asombro

https://doi.org/10.1017/langcog.2021.11 Published online by Cambridge University Press