



ORIGINAL ARTICLE

A cross-linguistic comparison of Spanish and English semantic norms: Looking at core features

Jorge Vivas^{1,*} , Boris Kogan² , Sofia Romanelli¹, Francisco Lizarralde¹ and Luis Corda¹

¹National University of Mar del Plata, Buenos Aires, Argentina and ²National University of Mar del Plata, Buenos Aires, Argentina, and National Agency of Scientific and Technological Promotion, Buenos Aires, Argentina

*Corresponding author. Email: jvivas@mdp.edu.ar

(Received 25 October 2018; revised 20 August 2019; accepted 18 September 2019; first published online 11 February 2020)

Abstract

It has been suggested that human communities that share their basic cultural foundations evince no remarkable differences concerning the characterization of core concepts. However, the small but existing differences among them reflect their sociocultural diversity. This study compares 219 concrete concepts common to both Spanish and English semantic feature norms in order to assess whether core features of concepts follow a universal or cultural language-specific pattern. Concepts were compared through a geometric technique of vector comparison in the Euclidean n -dimensional space alongside the calculation of the network's degree of centrality. The role of cognate status was also explored by repeating the former analysis separating cognate from noncognate words. Taken together, our data show that languages are structurally similar independent of the cognate status of words, further suggesting that there are some sort of core features common to both languages.

Keywords: comparison; conceptual representation; core features; cross-linguistic; feature norms

Does featural description follow a universal or a language-specific pattern? This has been an overlooked question in the literature (Kremer & Baroni, 2011). Comparative studies of English norms and other languages such as Italian and German (Kremer & Baroni, 2011) seem to indicate no relationship between feature production and language-specific characteristics. However, given the recent publication of the Spanish norms (Vivas, Vivas, Comesaña, García Coni, & Vorano, 2017), there are no comparative studies of this language and English yet which could support the aforementioned findings that indicate no outstanding differences across languages in relation to the characterization of core features. The main purpose of the present article is to make a cross-linguistic comparison of the structural organization of semantic feature productions between Spanish and English so as to discuss empirical information about the relationship between semantic

processing and different languages. In addition, focusing on bilingual semantic memory accounts that propose higher conceptual overlap for concrete and cognate words (de Groot, 1992; Van Hell & de Groot, 1998) compared to abstract and non-cognate lexical items, the role of cognate status was assessed.

Recognizing which are the main features underlying the semantic representation of a concept is key to the understanding of models of semantic memory. Recent distributional and feature-based models of semantic memory have studied the properties and organization of the semantic features that are involved in the representation of concepts (Baroni, Murphy, Barbu, & Poesio, 2010; Bolognesi, 2017; Cree, McNorgan, & McRae, 2006; Sartori, Polezzi, Mameli, & Lombardi, 2005; Taylor, Moss, & Tyler, 2007; Vigliocco, Meteyard, Andrews, & Koustas, 2009; Vigliocco, Vinson, Lewis, & Garrett, 2004). The fact that semantic features are key to identify a given concept is widely accepted (Cree *et al.*, 2006; Taylor *et al.*, 2007) and extensively acknowledged by different theoretical models (e.g., Cree & McRae, 2003; Humphreys & Forde, 2001; Mahon & Caramazza, 2009). Sartori *et al.* (2005) even assign an excluding role to this factor in the organization of conceptual knowledge.

Due to the growing interest in the study of the role of semantic features in the organization of concepts, numerous research groups of different countries and languages have analyzed and elaborated feature production norms (Buchanan, Holmes, Teasley, & Hutchinson, 2012; Buchanan, Valentine, & Maxwell, 2019; De Deyne *et al.*, 2008; Devereux, Tyler, Geertzen, & Randall, 2014; Lenci, Baroni, Cazzolli, & Marotta, 2013; McRae, Cree, Seidenberg, & McNorgan, 2005; Moldovan, Ferré, Demestre, & Sánchez-Casas, 2015; Montefinese, Ambrosini, Fairfield, & Mammarella, 2014; Ruts *et al.*, 2004; Stein & de Azevedo Gomes, 2009; Vinson & Vigliocco, 2008; Vivas *et al.*, 2017; Zannino *et al.*, 2006). Typically, in the collection of semantic features, participants are presented with a set of concepts and are asked to produce the features that they believe best describe these concepts, a task known as a feature-listing task. This collection of features constitutes a semantic register of produced semantic features of a particular linguistic community. In addition, semantic feature production norms also provide several qualitative and quantitative measures (e.g., Wu & Barsalou, 2009; taxonomy, imaginability, relevance, and significance) that are crucial for the design of controlled experimental stimuli. For example, in order to avoid cultural biases and construct stimuli lists matched in psycholinguistic variables, researchers may need normative values of all those variables regarding a particular linguistic community.

In contrast, empirically derived semantic feature production norms play also a highly important role in the elaboration and testing of theories about semantic memory. Devereux *et al.* (2014) suggest that theories about the representation and processing of concepts have improved from models based on information derived from the semantic production norms. This information refers not only to the identity of the features produced in the norms but also to their statistical properties. In this framework, statistical data about the frequency of occurrence and the probability of feature co-occurrence are proposed as the fundamental organizational principles of cognitive models of semantic memory. This data allows for the development of precise quantitative theories about the organization of conceptual knowledge to be explored by means of computational models (Al Farsi, 2018; Baroni & Lenci, 2010; Cree *et al.*, 2006; Lenci, 2018; Riordan & Jones, 2011). It is

worth mentioning that we agree with Riordan and Jones (2011) that featural and distributional models should not be conceptualized as competing theories; the focus should rather be on understanding the cognitive mechanisms human beings employ to integrate the two sources.

Gentner and Goldin-Meadow (2003) state that the dominant position within cognitive psychology in the last few decades has been that (a) conceptual structure is relatively constant in its core features across cultures, and (b) conceptual structure and semantic structure are closely coupled. Regarding the latter distinction, it is important to point out that “conceptual structure” refers to the mental representations of a given entity, while “semantic structure” is a wider term that encompasses the general organization of semantic memory, including various linguistic units and multimodal meanings assigned to words (Vigliocco & Filipovic, 2004). In this sense, the contribution of this study would be to provide a description of the different degrees of association between the paired productions of both languages in order to address the extensively discussed problem related to the effect of language on the structuring of concrete conceptual representations.

Against this background, we present a cross-linguistic comparison of Argentine Spanish and American English semantic norms for 219 concrete concepts (the intersection between both norms) in order to explore to what extent these norms reflect properties observed across languages (i.e., universal) and to what extent they are language specific. We expect concept descriptions to be generally stable across these language norms, providing some evidence that concrete words have conceptual cores.

Method

Stimuli

Stimuli were collected from the English (McRae et al., 2005) and Spanish (Vivas et al., 2017) sets of semantic feature norms. The English norms comprised 540 concepts and the Spanish one 400. For this study we selected 219 concepts common to both languages from 19 different concept categories.

Procedure

The translation of Spanish concepts to English was key to allow the matching up of concepts between languages. The English translation was made by a native Spanish speaker who lived in the United States for 10 years and by a native American English speaker who specialized in literature.

A very important aspect to consider when creating a semantic feature norm is to guarantee that synonymous features are registered in the same way within and between concepts. Given that semantic feature production norms are generated from the contribution of a large number of participants, and that each of them uses different words to characterize the concepts included in the worksheets they have to complete, it is remarkable to notice the different words spontaneously produced by them in order to refer to the same feature. For example, in order to characterize the Spanish concept “sol” (sun), some participants wrote “amarillo” (yellow) while

others “es amarillo” (it is yellow). Such cases, which are very frequent in the norms, are considered superficial variations of the same semantic content because there is no semantic difference among “amarillo” (yellow masculine gender), “amarillento” (yellowish), and “es de color amarillo” (it is yellow/it is yellow in color), but a difference in the way of referring to the same content. This variability should be reduced because of two fundamental reasons; first, the mass of data to process should become manageable, and second, the values of many of the variables that are crucial for semantic features (distinctiveness, relevance and production frequency, etc.) would be erroneously calculated if the spontaneous productions of semantic features remain unaltered. For example, let us consider how the variable frequency behaves for the Spanish concept “silla” (chair): 11 participants produced “tiene patas” (it has legs), 7 participants wrote “posee patas” (it possesses legs), and 6 simply responded “patas” (legs). If we consider these cases in isolation, these frequency values are not very high (taking into account that the highest value is 30, as each concept was characterized by 30 participants). However, if these responses are collapsed into a single feature (in this case, the resulting semantic feature is “tiene patas” [it has legs]), the frequency of production of the resulting semantic feature considerably increases, reaching a value of 24. In this study, we carried out a second interlinguistic coding (between the semantic features of both norms) for the same aforementioned reasons and under the same criteria. It is accepted a priori that the coding of responses may bring difficulties, as mentioned by Rogers *et al.* (2004), as researchers may vary in the way they code and analyze responses. However, as stated by these authors, this problem becomes less serious if it is assumed that verbally listed features do not constitute the conceptual representation per se. In this sense, the importance of the homogenization of data should not be underestimated because of the reasons already stated. What is very important then is to have a unified corpus with theoretically founded and empirically adequate coding criteria. These criteria have to be agreed on by the different groups of researchers devoted to the elaboration of semantic feature production norms. Once agreed upon, criteria should be rigorously applied to guarantee the replication of procedures.

Data analysis

The semantic production norms in English (McRae *et al.*, 2005) and Spanish (Vivas *et al.*, 2017) were elaborated on a basis of 540 and 400 concrete concepts, respectively, using an intersection between both norms of 219 concepts. For the generation of verbal descriptors and the verbal semantic field for the 219 concepts, the production frequency of features for each concept was considered. Based on that, the 438.txt files generated contained the name of the concept alongside its most frequent featural descriptions. To calculate the semantic relationship between concepts of both languages, we used the cosine between vectors, which could be considered analogous to correlation coefficients (the more semantically related two words are, the higher is its cosine value). Specifically, we used the standard n -dimensional inner product for the Euclidean space to compare vectors using the angle they form with each other. Parallel vectors represent the greatest similarity (i.e., a cosine value of 1) while orthogonal vectors represent the greatest difference (i.e., a cosine value

of 0; Kintsch, 2001). This calculation was done with the software SynonymFinder (Beta version 1.1; <http://iaai.fi.mdp.edu.ar:8080/sfweb>). With the result of the calculation of the cosine between each concept pair, the software generates a square matrix model, which is the one in which columns and rows refer to the same set of entities (see Matrix 1, in the online-only supplementary material). Based on this matrix, we calculated how many concepts did or did not have their strongest link (in terms of cosine values) with their equivalent in the other language. Then, we calculated the mean of those values for both the concepts that did and did not have their strongest link with their equivalent in the other language, also looking for intralinguistic cases that showed similar values to those observed in interlinguistic examples. Specifically, for the concepts that did not coincide with their equivalent in the other language, we calculated the difference in terms of cosine values between each concept's highest interlinguistic associate and its equivalent in the other language. Finally, Johnson's method (1967) was applied to do the hierarchical clustering by using UCINET software (Borgatti, Everett, & Freeman, 2002), taking as a starting point this matrix. We repeated this analysis for both cognate and noncognate words.

The results of the clustering analysis was plotted using the social network visualization software NetDraw (Borgatti et al., 2002). In addition, two procedures were followed: (a) an analysis of the association level between the English and the Spanish matrices by means of the quadratic assignment procedure (QAP), which is principally used to test the association between networks in Social Network Analysis (Borgatti et al., 2002), and (b) the calculation of the degree centrality of the concepts from both matrices so as to analyze the degree of association between concepts from both languages. The latter analysis was conducted in SPSS 19. The QAP procedure was used here to compute the correlation between two square matrices, English and Spanish in this case. The algorithm proceeds in two steps. In the first one, it computes Pearson's correlation coefficient between corresponding cells of the two data matrices. In the second step, it randomly permutes rows and columns synchronously of one matrix and recomputes the correlation and other measures. This step is carried out thousands of times in order to compute the proportion of times that a random measure is larger than or equal to the observed measure calculated in Step 1. A low proportion ($<.05$) suggests a strong relationship between the matrices that is unlikely to have occurred randomly. The centrality degree calculates the degree and normalized degree centrality of each entity. The number of entities adjacent to a given entity in a symmetric matrix is the degree of that entity (Wasserman & Faust, 1998).

Results

The vector comparison method showed that the average of the strongest link between a concept and its equivalent in the other language had a value of .7 in 87.21% of the cases (being 0 and 1 indicators of a null and a full relationship, respectively), while the remaining 12.79% had an average value of .5. In addition, and regarding concepts that did not coincide with its equivalent in the other language, differences between each concept's highest associate and its equivalent in the other

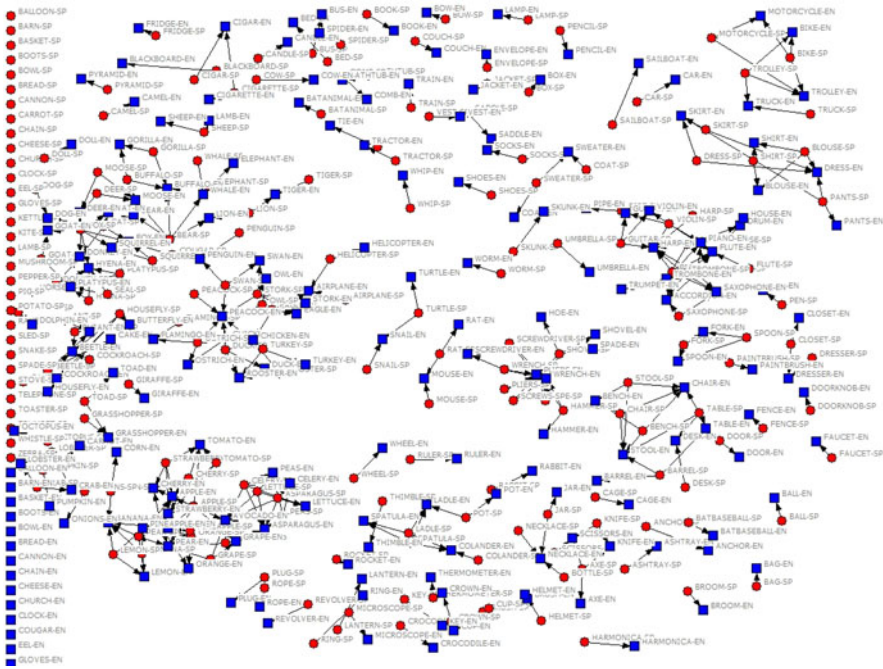


Figure 2. Cluster diagram showing the links (in terms of cosine values) between concepts of the English (blue squares) and Spanish (red dots) semantic norms in our database. This diagram was plotted with NetDraw Software.

A Fisher's transformation test (Fisher, 1915) revealed that the difference between the correlation coefficients of cognates and noncognates was not significant ($Z = 1.33$, $p = .0918$). Typically, cognates are defined as translation equivalents that overlap in meaning and lexical form (Prior, MacWhinney, & Kroll, 2007). In our study, the criterion used to classify words into cognates and noncognates was that they shared at least 40% of phonological overlap (e.g., barrel–barril, tomato–tomate). Furthermore, as can be seen in the hierarchical clustering (see Figure 2), concepts that share the same category are grouped together. Again, the same was found for both cognates (see Figure 3) and noncognates (see Figure 4).

Finally, due to the nonparametric distribution of cases, a Spearman correlation analysis of the degree of centrality of both networks was conducted. This represents an indirect measure of the amount of semantic features that each concept shares with the rest in the network. Results showed a strong positive correlation between the degree of centrality of both Spanish and English concepts that comprised the networks ($r = .726$, $n = 219$, $p = .001$).

Discussion

This work performed a cross-linguistic comparison of Spanish (Vivas et al., 2017) and English (McRae et al., 2005) semantic feature production norms comprising

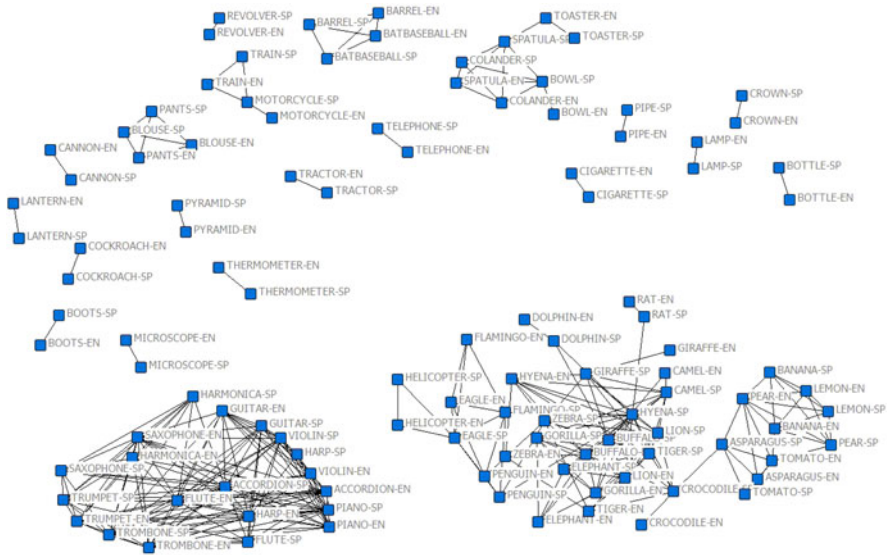


Figure 3. Cluster diagram showing the links (in terms of cosine values) between cognate concepts of the English (-EN) and Spanish (-SP) semantic norms in our database. This diagram was plotted with NetDraw Software.

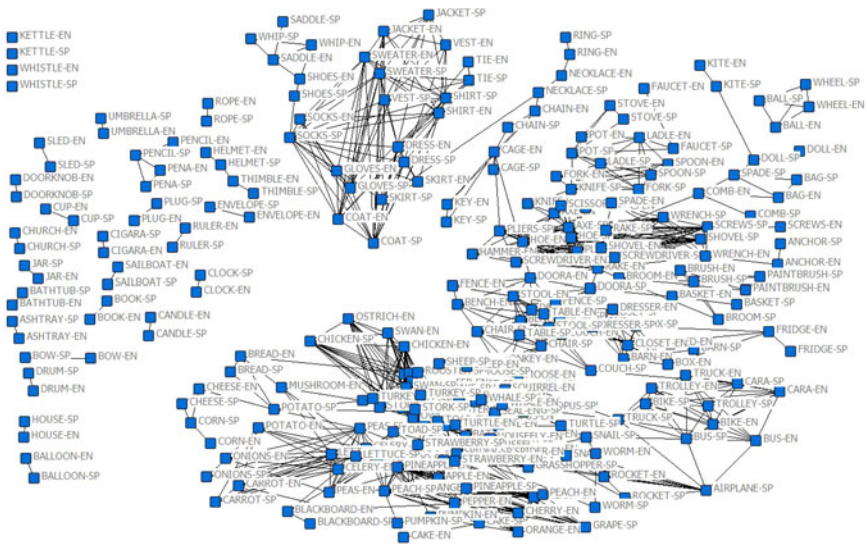


Figure 4. Cluster diagram showing the links between noncognate concepts of the English (-EN) and Spanish (-SP) semantic norms in our database. This diagram was plotted with NetDraw Software.

219 common concrete concepts. The geometric technique of vector comparison in the Euclidean n -dimensional space was used alongside the calculation of the network's degree of centrality. We found that conceptual structure was similar in both languages independent of the cognate status of words, further suggesting

the existence of some sort of core features across languages. Below we address these issues in turn.

Before discussing the results obtained in the present study, it is worth explaining what we mean by “coincidence with its equivalent.” Under this term, we include those cases in which a concept obtains its strongest link with its equivalent in the other language, that is to say, interlinguistic cases. Then, for example, “ashtray” from the English norm has its strongest interlinguistic link (value .7) with “cenicero” (ashtray) from the Spanish norm. In these cases, the strongest link, which never reaches a “value of 1” but averages values higher than .6, was found for the relationship with its equivalent in the other language in 87.21% of the cases.

Murphy (2002, p.1) defines a concept as “a mental representation corresponding to that category (the class of objects in the world).” Moreover, he states that concepts “are a kind of mental glue” as they connect our previous experiences to our present and future interactions with objects. In this sense, we consider that the 87.21% of the interlinguistic coincidence found in our study gives support to the hypothesis we have put forward: as a concept-generation process, feature production norms would have at least some kind of universal core, independent of the language that uses it. The aforementioned hypothesis aligns with what Gentner and Goldin-Meadow (2003) have stated in relation to the dominant position within cognitive psychology in recent years, that (a) conceptual structure is relatively constant in its core features across cultures, and (b) conceptual structure and semantic structure are closely coupled.

We found a high percentage of coincidence between the cosine values of the English and Spanish concept equivalents. In addition, the QAP analysis showed a high correlation between both language networks even for both cognates and noncognates, further suggesting that this relationship is independent of the cognate status of words. Finally, though tentatively, we also observed some intralinguistic examples of concepts that exhibited similar values to the ones yielded in the interlinguistic analysis. For instance, the relationship between the cosine values of the words “ant” (hormiga) and “beetle” (escarabajo) is quite similar both inter- and intralinguistically. This information could be indicative of some sort of intra- and interlinguistic conceptual network (Lamas, Vivas, & Vorano, 2012; Vorano, Zapico, Corda, Vivas, & Vivas, 2014), and could coincide with the proposal by Moss, Tyler, and Taylor (2009) when they state that in the framework of these models, concepts are represented as activation patterns of several nodes within a neural network; in other words, these nodes correspond to various semantic features.

As shown in the clusterization, concepts are grouped together by type of related category (musical instruments, things that fly, insects, fruits/vegetables, etc.), and this phenomenon is also evident with concepts that do not coincide with its equivalent in the other language. The same was found in the degree of centrality analysis that was carried out. Specifically, when degree of centrality values were organized in descending order, a clear distinction between living and nonliving things was observed, representing the highest and lowest values, respectively. Moreover, it is important to notice that within living things, the animal category was the one that shared the greatest number of attributes with the rest of the members in the network. This could also be considered another strong indicator not only for the existence of a conceptual core but also for the presence of another level of intra- and interlinguistic association that is not so specific but still constant.

These results are consistent with what Kremer and Baroni (2011) proposed, as they state that when considering language communities that share similar cultural foundations, research studies suggest that there are no remarkable differences across languages in relation to the semantic features used by native speakers to characterize concepts. In this sense, it is worth mentioning that these results may change dramatically if the languages compared entailed significant cultural differences. More interesting, the present study found few linguistic as well as cultural differences worth mentioning. For instance, one of these differences in our study is evident in the concept “pavo” (turkey) from the Spanish norm (Vivas *et al.*, 2017) that coincides in its higher value with the concept “peacock” from the English norm (McRae *et al.*, 2005). The first cultural difference is evidenced by the fact that “turkey” is a typical everyday food in the United States, and a traditional Thanksgiving and Christmas dish. However, it is rarely eaten by Argentinians, from whom Spanish semantic features were collected. This difference generates variation in different characteristics of the conceptual structure, like “familiarity.” Second, both linguistic labels in English differ phonologically, “turkey” and “peacock,” which does not allow speakers to make associations at a phonological level. Conversely, the Spanish concepts “pavo” (turkey) and “pavo real” (peacock) are homonymous, with the exception that the latter concept is adjectivized, which would imply the access to a phonological level of association before other levels.

In contrast, there are 12.79% of concepts whose strongest link does not coincide with the equivalent in the other language. We believe, however, that this lack of coincidence does not contradict what we have claimed earlier about the existence of core features. As illustrated in Figure 1, for the Spanish concept “lechuga” (lettuce), the English equivalent, “lettuce,” occupies the third position, with a relationship value between each other of .745. This could be interpreted as saying that its semantic distance in relation to the Spanish concept “lechuga” (lettuce) is larger than the distance between “asparagus” in English and “celery” in English, with values of .785 and .766, respectively. Even though this result is correct, it is surprising that the difference between the highest value for “asparagus” and “lettuce,” which would be the corresponding one but occupies the third position, is only .04. If the column in Figure 1 labeled “difference” is analyzed, even though the maximum value for the concept “tornillo” (screws) is high (.377), in general, the differences were not very significant, reaching an average value of .11090, with a standard deviation of 0.094080, which would support our claim. Moreover, if the total values of cases per position are observed, it would be noted that the larger concentration is found in the second and third positions with 76% of the total, and that values decrease significantly and abruptly from the fourth column on. Once more, we consider this result to be indicative of the existence of strong interlinguistic relationships between concepts, even in cases where there is no coincidence. Finally, it is crucial to acknowledge that the present results may only apply to concrete concepts. As shown by some studies, abstract concepts might lack of conceptual core (Pulvermüller, 2018) and exhibit remarkable processing differences (Bolognesi & Sten, 2018) compared to concrete concepts.

To sum up, and in line with the findings reported by Kremer and Baroni (2011) about the relationships between the semantic production norms in English, German, and Italian, the results of the present study show a general consistency

in the distribution of patterns in English and Spanish. The 87.21% of the concepts analyzed in this study revealed the strongest interconceptual link with interlinguistic pairs with an average value of .7, being 0 and 1 indicators of a null and a full relationship, respectively.

Conclusion

A set of 219 concrete concepts common to both English and Spanish semantic feature norms was compared to assess whether concepts' core features evidence universal or language-specific properties. Even though few linguistic as well as cultural differences were observed affecting the structuring of conceptual representation of concrete words, our results suggested that there are no remarkable differences across languages in terms of the semantic features used by native speakers to describe and characterize concrete concepts. In this sense, our findings are indicative of the existence of some sort of core features common to both languages.

Acknowledgment. This research was supported by National Agency for Scientific and Technological Promotion Grant PICT 2015-0983 (to J.V.).

References

- Al Farsi, B. (2018). Word meaning in word identification during reading: Co-occurrence-based semantic neighborhood density effects. *Applied Psycholinguistics*, *39*, 779–809. doi: [10.1017/S0142716417000583](https://doi.org/10.1017/S0142716417000583)
- Baroni, M., & Lenci, A. (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, *36*, 673–721. doi: [10.1162/coli_a_00016](https://doi.org/10.1162/coli_a_00016)
- Baroni, M., Murphy, B., Barbu, E., & Poesio, M. (2010). Strudel: A corpus-based semantic model based on properties and types. *Cognitive Science*, *34*, 222–254.
- Bolognesi, M. (2017). Using semantic feature norms to investigate how the visual and verbal modes afford metaphor construction and expression. *Language and Cognition*, *9*, 525–552. doi: [10.1017/langcog.2016.27](https://doi.org/10.1017/langcog.2016.27)
- Bolognesi, M., & Steen, G. (Eds.) (2018). Abstract concepts: Structure, processing and modeling. *TopiCS in Cognitive Science*, *10*.
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). Ucinet for Windows: Software for social network analysis [Computer software]. Harvard, MA: Analytic Technologies.
- Buchanan, E. M., Holmes, J. L., Teasley, M. L., & Hutchison, K. A. (2012). English semantic word-pair norms and a searchable Web portal for experimental stimulus creation. *Behavior Research Methods*, *45*, 746–757. doi: [10.3758/s13428-012-0284-z](https://doi.org/10.3758/s13428-012-0284-z)
- Buchanan, E. M., Valentine, K. D., & Maxwell, N. (2019). English semantic feature production norms: An extended database of 4,436 concepts. *Behavior Research Methods*. Advance online publication.
- Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *JEP: Learning, Memory, and Cognition*, *32*, 643.
- Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163–201. doi: [10.1037/0096-3445.132.2.163](https://doi.org/10.1037/0096-3445.132.2.163)
- De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., & Storms, G. (2008). Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behavior Research Methods*, *40*, 1030–1048. doi: [10.3758/BRM.40.4.1030](https://doi.org/10.3758/BRM.40.4.1030)
- de Groot, A. M. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1001–1018. doi: [10.1037/0278-7393.18.5.1001](https://doi.org/10.1037/0278-7393.18.5.1001)
- Devereux, B. J., Tyler, L. K., Geertzen, J., & Randall, B. (2014). The Centre for Speech, Language and the Brain (CSLB) concept property norms. *Behavior Research Methods*, *46*, 1119–1127. doi: [10.3758/s13428-013-0420-4](https://doi.org/10.3758/s13428-013-0420-4)

- Fisher, R. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, *10*, 507–521. doi: [10.2307/2331838](https://doi.org/10.2307/2331838)
- Gentner, D., & Goldin-Meadow, S. (2003). Whither whorf. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition*. Cambridge, MA: MIT Press.
- Humphreys, G. W., & Forde, E. (2001). Category specificity in mind and brain. *Behavioral and Brain Sciences*, *24*, 497–504.
- Johnson, S. (1967). Hierarchical clustering schemes. *Psychometrika*, *32*, 241–253.
- Kintsch, W. (2001). Predication. *Cognitive Science*, *25*, 173–202.
- Kremer, G., & Baroni, M. (2011). A set of semantic norms for German and Italian. *Behavior Research Methods*, *43*, 97–109. doi: [10.3758/s13428-010-0028-x](https://doi.org/10.3758/s13428-010-0028-x)
- Lamas, V., Vivas, J., & Vorano, A. (2012). Comparación de atributos semánticos entre diferentes lenguas. In A. Vigaro (Ed.), *IV Congreso Internacional de Investigación en Psicología*. Buenos Aires, Argentina: Universidad Nacional de La Plata, Facultad de Psicología.
- Lenci, A. (2018). Distributional models of word meaning. *Annual Review of Linguistics*, *4*, 151–171. doi: [10.1146/annurev-linguistics-030514-125254](https://doi.org/10.1146/annurev-linguistics-030514-125254)
- Lenci, A., Baroni, M., Cazzolli, G., & Marotta, G. (2013). BLIND: A set of semantic feature norms from the congenitally blind. *Behavior Research Methods*, *45*, 1218–1233. doi: [10.3758/s13428-013-0323-4](https://doi.org/10.3758/s13428-013-0323-4)
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review Psychology*, *60*, 27–51. doi: [10.1146/annurev.psych.60.110707.163532](https://doi.org/10.1146/annurev.psych.60.110707.163532)
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*, 547–559. doi: [10.3758/BF03192726](https://doi.org/10.3758/BF03192726)
- Moldovan, C., Ferré, P., Demestre, J., & Sánchez-Casas, R. (2015). Semantic similarity: Normative ratings for 185 Spanish noun triplets. *Behavior Research Methods*, *47*, 788–799. doi: [10.3758/s13428-014-0501-z](https://doi.org/10.3758/s13428-014-0501-z)
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods*, *46*, 887–903. doi: [10.3758/s13428-013-0405-3](https://doi.org/10.3758/s13428-013-0405-3)
- Moss, H. E., Tyler, L. K., & Taylor, K. I. (2009). Conceptual structure. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics*. Oxford: Oxford University Press.
- Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Prior, A., MacWhinney, B., & Kroll, J. F. (2007). Translation norms for English and Spanish: The role of lexical variables, word class, and L2 proficiency in negotiating translation ambiguity. *Behavior Research Methods*, *39*, 1029–1038.
- Pulvermüller, F. (2018). Neurobiological mechanisms for semantic feature extraction and conceptual flexibility. *Topics in Cognitive Science*, *10*, 590–620.
- Riordan, B., & Jones, M. (2011). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science*, *3*, 303–345. doi: [10.1111/j.1756-8765.2010.01111.x](https://doi.org/10.1111/j.1756-8765.2010.01111.x)
- Rogers, T. T., Garrad, P., McClelland, J. L., Lambon Ralph, M. A., Bozeat, S., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*, 205–235.
- Ruts, W., De Deyne, S., Ameel, E., Vanpaemel, W., Verbeemen, T., & Storms, G. (2004). Dutch norm data for 13 semantic categories and 338 exemplars. *Behavior Research Methods, Instruments, & Computers*, *36*, 506–515.
- Sartori, G., Polezzi, D., Mameli, F., & Lombardi, L. (2005). Feature type effects in semantic memory: An event related potentials study. *Neuroscience Letters*, *390*, 139–144. doi: [10.1016/j.neulet.2005.08.015](https://doi.org/10.1016/j.neulet.2005.08.015)
- Stein, L. M., & de Azevedo Gomes, C. F. (2009). Normas brasileiras para listas de palavras associadas: Associação semântica, concretude, frequência e emocionalidade. *Psicologia: Teoria e Pesquisa*, *25*, 537–546.
- Taylor, K. I., Moss, H. E., & Tyler, L. K. (2007). The conceptual structure account: A cognitive model of semantic memory and its neural instantiation. In J. Hart & M. Kraut (Eds.), *The neural basis of semantic memory* (pp. 265–301). Cambridge: Cambridge University Press.

- Van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*, 193–211. doi: [10.1017/S1366728998000352](https://doi.org/10.1017/S1366728998000352)
- Vigliocco, G., & Filipovic, L. (2004). *From mind in the mouth to language in the mind*. Cambridge, MA: MIT Press.
- Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*, 219–247. doi: [10.1515/LANGCOG.2009.011](https://doi.org/10.1515/LANGCOG.2009.011)
- Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*, 422–488. doi: [10.1016/j.cogpsych.2003.09.001](https://doi.org/10.1016/j.cogpsych.2003.09.001)
- Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*, 183–190. doi: [10.3758/BRM.40.1.183](https://doi.org/10.3758/BRM.40.1.183)
- Vivas, J., Vivas, L., Comesaña, A., García Coni, A., & Vorano, A. (2017). Spanish semantic feature production norms for 400 concrete concepts. *Behavior Research Methods*, *49*, 1095–1106. doi: [10.3758/s13428-016-0777-2](https://doi.org/10.3758/s13428-016-0777-2)
- Vorano, A., Zapico, G., Corda, L., Vivas, J., & Vivas, L. (2014). Comparación de atributos semánticos entre castellano rioplatense e inglés. *VI Congreso Marplatense de Psicología*. Mar del Plata, Argentina: Universidad Nacional de Mar del Plata, Facultad de Psicología.
- Wasserman, S., & Faust, K. (1998). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
- Wu, L.-L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, *132*, 173–189.
- Zannino, G. D., Perri, R., Pasqualetti, P., Di Paola, M., Caltagirone, C., & Carlesimo, G. (2006). *Neuropsychologia*, *44*, 1017–1028. doi: [10.1016/j.neuropsychologia.2005.11.006](https://doi.org/10.1016/j.neuropsychologia.2005.11.006)

Cite this article: Vivas, J., Kogan, B., Romanelli, S., Lizarralde, F., and Corda, L. (2020). A cross-linguistic comparison of Spanish and English semantic norms: Looking at core features. *Applied Psycholinguistics* *41*, 285–297. <https://doi.org/10.1017/S0142716419000523>