Are valence and arousal related to the development of amodal representations of words? A computational study

1. José Ángel Martínez-Huertas (Department of Methodology of Behavioral Sciences; Universidad Nacional de Educación a Distancia) – jamartinez@psi.uned.es.
2. Guillermo Jorge-Botana* (Department of Psychobiology and Methodology of Behavioral Sciences; Universidad Complutense de Madrid) - guijorge@ucm.es.
3. Alejandro Martínez-Mingo (Department of Social Psychology and Methodology; Universidad Autónoma de Madrid) - a.martinez.m90@gmail.com.
4. Diego Iglesias (Department of Social Psychology and Methodology; Universidad Autónoma de Madrid) - diego.iglesiaso@estudiante.uam.es.
5. Ricardo Olmos (Department of Social Psychology and Methodology; Universidad Autónoma de Madrid) - ricardo.olmos@uam.es.

Cognition and Emotion

This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available, upon publication, via its DOI: https://doi.org/10.1080/02699931.2023.2283882.

* = Corresponding author:

Guillermo Jorge-Botana, Department of Psychobiology and Methodology of Behavioral Sciences, Faculty of Psychology, Campus de Somosaguas, Universidad Complutense de Madrid, Pozuelo de Alarcón, Madrid, 28049. Madrid, Spain. E-mail: guijorge@ucm.es.

ABSTRACT

In this study, we analyzed the relationship between the amodal (semantic) development of words and two popular emotional norms (emotional valence and arousal) in English and Spanish languages. To do so, we combined the strengths of semantics from vector space models (vector length, semantic diversity, and word maturity measures), and feature-based models of emotions. First, we generated a common vector space representing the meaning of words at different developmental stages (five and four developmental stages for English and Spanish, respectively) using the Word Maturity methodology to align different vector spaces. Second, we analyzed the amodal development of words through mixed-effects models with crossed random effects for words and variables using a continuous time metric. Third, the emotional norms were included as covariates in the statistical models. We evaluated more than 23,000 words, whose emotional norms were available for more than 10,000 words, in each language separately. Results showed a curve of amodal development with an increasing linear effect and a small quadratic deceleration. A relevant influence on the amodal development of words was found only for emotional valence (not for arousal), suggesting that positive words have an earlier amodal development and a less pronounced semantic change across early lifespan.

Keywords: development; words; emotions; emotional valence; semantics.

INTRODUCTION

It is known that language plays an important mechanism in emotion perception and emotion understanding from infancy through older adulthood. For example, authors advocating for the *psychological constructionist hypothesis* affirm that language would support the acquisition of emotional knowledge, which would influence emotional experiences across life span (Shablack & Lindquist, 2019). Among other experimental procedures, multiple studies have been conducted to study human emotional knowledge using emotional words or concepts, that is, evaluating the emotional properties of linguistic information (e.g., Hoemann & Feldman-Barrett, 2019; see also the review of Shablack & Lindquist, 2019). In this way, some authors defend that there are interactions between emotions and cognition that influence various aspects of human behavior, including language development (Panksepp, 2007; Pexman, 2019). On the contrary, it is not still clear how the meaning of words, that is, their amodal[1] (semantic) representations, change across life span and how emotional norms, like emotional valence or arousal, influence such amodal development. Thus, the present study aims to analyze how emotional valence and arousal influence the development of amodal representations of words using a longitudinal computational study in two different languages (English and Spanish). To do so, we combine the strengths of semantics from vector space models (e.g., Günther et al., 2019; Jorge-Botana et al., 2020) and feature-based models of emotions (e.g., Fraga et al., 2018).

[1] Amodal representations of words can be understood as semantic representations. Amodal representations refer to how words are related to each other propositionally and conceptually (e.g., "pain" and "blood" are semantically related by means of common amodal contexts). In contrast, modal representations refer to encoded sensorimotor and emotional information in a purer state (e.g., "pain" and "blood" encode negative emotional information).

Previous research found that some parts of the amodal representations of words were useful for propagating the emotionality of words (Hoffman et al., 2018), although there was no isomorphism between emotional and modal representations of words (see the *specific dimensionality hypothesis*; Martínez-Huertas, Jorge-Botana, Luzón, & Olmos, 2021). Other studies also found that such linking mechanisms might be instantiated during earlier developmental stages and could partially explain the emotionality of words at later adult developmental stages (e.g., Martínez-Huertas, Jorge-Botana, & Olmos, 2021). Among other differences, it was found that the propagation of emotions through language was more efficient in words with moderately early amodal maturation (not necessarily finished) when their amodal representations were more semantically defined in the adult stage. That is, the emotionality of adult words can be better predicted from its amodal representations when the latter are more defined in the adult stage and are more similar to their respective child representations (Martínez-Huertas, Jorge-Botana, & Olmos, 2021). Thus, there might be a differential amodal development of words, which could have complex and bidirectional relationships with their emotional properties. Nonetheless, a limitation of previous research is that the amodal development of words was modeled using a discrete time representation with only two time points (i.e., child and adult developmental stages; e.g., Martínez-Huertas, Jorge-Botana, & Olmos, 2021). The present study aims to hone the modeling of the amodal development of words by capturing the continuous-time dynamical amodal changes of words.

One relevant finding about human development for our objective is that there is a shift in children's emotional processing from anchoring on valence information (positive vs. negative) to multidimensional emotions across development (e.g., Hoemann &

Feldman-Barrett, 2019; Nook et al., 2017). This means that there would be a bidirectional, complex, and complementary learning of language and emotional information during the life span. Thus, to fully understand the language-emotion relationship during development, it is necessary to have evidence of the kind of emotional words that are learned earlier and how their emotionality could shape the amodal development of their amodal meanings. Considering previous empirical findings, we follow the *circumplex model of affect* (Russell, 1980), concretely the valence and arousal dimensions, as an operative definition of the emotionality of words during early development.

METHODS

Longitudinal computational models

Different vector spaces were generated for each developmental stage in English and Spanish, separately. The English vector spaces were generated using the *Touchstone Applied Science Associates* (TASA) corpus (Ivens & Koslin, 1991; Landauer et al., 1998; http://lsa.colorado.edu/spaces.html), which is a popular text corpus in psychological literature. The Spanish vector spaces were generated using a Spanish corpus about childish fables and tales for the early developmental stages (Jorge-Botana et al., 2017) and the *LEXESP* corpus for the adult developmental stage (Sebastián et al., 2000). Table 1 presents the properties of English and Spanish linguistic corpora[2] used to generate the vector spaces of each developmental stage. The TASA corpus has 37,650 documents with 23,057 unique

[2] Note that we use linguistic corpora of academic texts, fables, and tales adapted to each developmental stage, although a more valid and desirable scenario would have been to use natural language corpora produced by children at different developmental stages.

words in the following developmental stages: 8-to-9 years, 11-to-12, 14-to-15, 17-to-18, and adult. The Spanish corpus has 263,928 paragraphs and 25,413 unique words in the following developmental stages: 0-to-9 years, 9- to-12, 12-to-16, and adult. Note that the total number of documents and unique words of each language is referred to the adult developmental stage linguistic corpora (that is, given that they were cumulative corpora, previous developmental stages have a smaller number of documents and unique words). Text corpora of each developmental stage in each language were processed with standard procedures of the LSA technique (e.g., Landauer & Dumais, 1997) using *Gallito Studio* software (Jorge-Botana et al., 2013) with the log-entropy weighted function. After *Singular Value Decomposition* (SVD), 300 dimensions were imposed for the vector spaces.

TABLE 1. Properties of English and Spanish linguistic corpora used to generate the vector spaces of each developmental stage.

English linguistic corpora			**Spanish linguistic corpora**		
Developmental stage	**Documents**	**Words (unique terms)**	**Developmental stage**	**Documents**	**Words (unique terms)**
8-to-9 years	6,974	6,304	0-to-9 years	31,722	7,286
11-to-12 years	17,949	12,962	9-to-12 years	59,746	10,507
14-to-15 years	22,211	15,340	12-to-16 years	75,434	13,209
17-to-18 years	28,882	18,804	Adult (21 years)	263,928	25,413
Adult (21 years)	37,650	23,057			

Note: Only words that appeared in more than five documents were used to generate the vector space model.

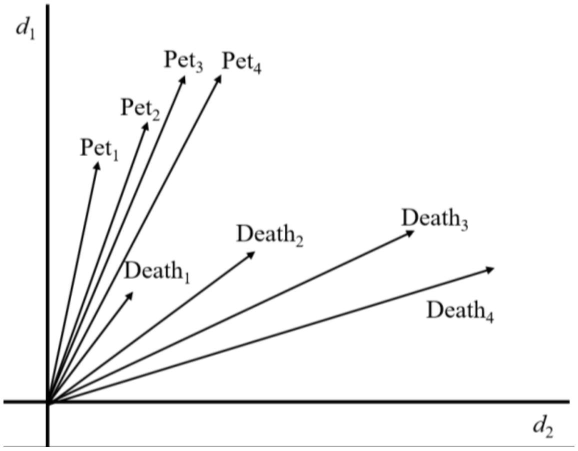
The semantic spaces were aligned at each developmental stage using the *Word Maturity* technique (Kireyev & Landauer, 2011; Landauer et al., 2011; Jorge-Botana et al., 2017, 2018). Basically, this technique is a mathematical procedure that rotates the matrices of two vector spaces by means of the centering of both spaces, establishing a common scaling, and applying a rotation matrix to make them homologous. This rotation matrix is obtained from the minimization of the Frobenius norm:

$$||XQ - Y|| \rightarrow \min \qquad [1]$$

where X and Y are the centered and scaled vector spaces and Q is the rotation matrix to be found (see Jorge-Botana et al., 2018 for a full explanation). This alignment was repeated four times for the English vector space as it has five developmental stages, and three times for the Spanish vector space as it has four developmental stages. This alignment of vector spaces was used to make vector representations from all the developmental stages comparable. Figure 1 presents a graphical illustration of the hypothetical vector representation of two words at four time points in a fictional vector space with two dimensions to ease its interpretability. Being the vector representations projected on the same basis, words have independent vector representations in each developmental stage, if they exist (e.g., there are words that might not be acquired until later developmental stages, and words only have vector representations when they are available at that time point). Then, measures as distances between vectors and vector lengths can be extracted from the independent vector representations of the same word at different developmental stages.

Figure 1

Graphical illustration of the vector representations of two words (Pet, Death) at four time points (t_1, t_2, t_3, t_4) in a fictional bidimensional vector space.



Computational variables

Once all the vector spaces were aligned, different semantic measures were extracted for each word as indicators of its amodal development and maturation in each developmental stage for both languages. Concretely, we extracted vector length, semantic diversity, and word maturity using *Gallito Studio* software (Jorge-Botana et al., 2013). Vector length, $|\vec{v}_{ti}|$, is a measure of the quantity of information encoded in the vector of word *i* at time *t*. Semantic diversity, $SemD_{ti}$, is a measure of the degree to which different contexts associated with word *i* at time *t* vary in their meanings, which quantifies the context-dependent variations of word meaning (e.g., Hoffman et al., 2013). We computed semantic diversity as $1 - SemD_{ti}$ to capture the focalization of the semantics. Word maturity, WM_{ti}, is a measure of how mature a word *i* at time *t* with respect to its adult state is (Jorge-Botana et al., 2017, 2018). These measures were used to estimate a general state of amodal development of words using mixed-effects models with crossed random effects (MEMs-CR).

Emotional norms (valence and arousal)

The norms of emotional valence and arousal were selected from two public norm data sets in English (*N* = 13,915 words; Warriner et al., 2013) and Spanish (*N* = 14,031 words; Stadthagen-Gonzalez et al., 2017). Given that not all the words had a vector representation in the computational model, we used the emotional norms of 11,071 and 10,613 words in English and Spanish, respectively.

Multivariate longitudinal data analysis

Different MEMs-CR for words and computational variables were fitted using continuous time metrics in R software[3]. We followed Martínez-Huertas & Ferrer (2023) proposal, which can be adapted to the multivariate longitudinal analysis with linear and quadratic effects of time of this study using the following equation:

$$Y_{ivt} = \mu_{0iv} + \mu_{1iv} \cdot time_{t} + \mu_{2iv} \cdot time^{2}_{t} + e_{ivt}, \text{ with}$$

$$\mu_{0iv} = \mu_{0i} + \mu_{0v} + e_{0vt},$$

$$\mu_{1iv} = \mu_{1i} + \mu_{1v} + e_{1vt}, \text{ and}$$

$$\mu_{2iv} = \mu_{2i} + \mu_{2v} + e_{2vt} \qquad [2]$$

[3] The interpolation of values (that is, inferring trajectories of meaning change from the beginning of their acquisition to the adult state) is one of the objectives of *Word Maturity* methodology (e.g., Jorge-Botana et al., 2017, 2018). In this line, MEMs-CR with a continuous time metric can be an affordable tool to interpolate the complete trajectories of the amodal development of words for any time *t*, and to study the influence of covariates in such trajectories. The OSF project contains the R scripts to replicate the analyses.

where Y_{ivt} is the variable (computational measure) *v* for word *i* at time *t*, μ_{0iv} is the mean intercept of the sample for all variables, μ_{1iv} is the general linear slope of time, and μ_{2iv} is the general quadratic slope of time. These effects can vary for individual words and variables. Such sources of variability are captured in the random effects (being $\sigma^{(2)}{}_{0i}$ and $\sigma^{(2)}{}_{0v}$ the variances of the intercepts for words and variables, respectively; $\sigma^{(2)}{}_{1i}$ and $\sigma^{(2)}{}_{1v}$ the variances of the linear slopes for words and variables, in that order; and $\sigma^{(2)}{}_{2i}$ and $\sigma^{(2)}{}_{2v}$ the variances of the quadratic slopes for words and variables, respectively). Then, each y_{0iv} is a function of the random variance of the intercept of individual *i* and variable *v*, while y_{1iv} is a function of the random variance of the linear slope of individual *i* and variable *v*, and y_{2iv} is a function of the random variance of the quadratic slope of individual *i* and variable *v*, being the error term (e_{ivt}) dependent on the word, the variable, and the time.

First, different MEMs-CR were estimated to describe the general longitudinal trajectory of words and the aforementioned variables. To do so, we compared a null model, a model with linear effects of time, and a model with linear and quadratic effects of time, to describe longitudinal trajectories of amodal development. Second, different random structures of the selected MEM-CR were compared to determine the complexity of the random effects of the model. Likelihood ratio tests and relative fit indices were used to select the best-fitting model using a bottom-up strategy (that is, from the simpler to the most complex model; for a more exhaustive explanation about model selection in MEMs-CR, see Martínez-Huertas et al., 2022). Third, the standardized scores of the emotional norms (emotional valence and arousal) were introduced in the selected model as covariates. These analyses were conducted for both English and Spanish languages separately using

REML estimator. Variables were standardized with respect to the means and standard deviations of the first developmental stage in each language.

RESULTS

In the case of English words, it was found that a linear+quadratic effect of time was slightly more appropriate than just a linear effect ($\Delta\text{AIC} = 79, \Delta\text{BIC} = 69$), and the latter obtained a better model fit than the null model ($\Delta\text{AIC} = 25408, \Delta\text{BIC} = 25398$). The same pattern of results was found for Spanish words, where including the linear+quadratic effect of time obtained a better model fit than just the linear effect ($\Delta\text{AIC} = 972, \Delta\text{BIC} = 962$), and its fit was better than the null model ($\Delta\text{AIC} = 6776, \Delta\text{BIC} = 6766$). Additionally, it was found that including the random slopes of time effects for words presented a better model fit than just the random intercepts in both languages (English: $\chi^2(5) = 10717.0, p < .001, \Delta\text{AIC} = 1727, \Delta\text{BIC} = 1779$; Spanish: $\chi^2(5) = 16258.91, p < .001, \Delta\text{AIC} = 16269, \Delta\text{BIC} = 16318$), while the inclusion of random slopes for variables presented convergence problems. Thus, we selected a MEM-CR with linear+quadratic time effects with random intercepts for words and variables, and random slopes for words, to analyze the data of both languages.

Table 2 presents the results of the MEMs-CR that define the trajectories of amodal development for English and Spanish words. The model includes emotional valence and arousal as covariates. Then, the reported time effects can be understood as the trajectories of words with mean scores of valence and arousal[4]. In both languages, it is possible to interpret an important linear effect of time (μ_{1iv}=.159, *SE*=.007, *t*=22.525, *p*<.001 for English words; and μ_{1iv}=.190, *SE*=.007, *t*=26.351, *p*<.001 for Spanish words) and a smaller negative quadratic effect (μ_{2iv}=-.001, *SE*=.000, *t*=-3.911, *p*<.001 for English words; and μ_{2iv}=-.004, *SE*=.000, *t*=-17.346, *p*<.001 for Spanish words) that sightly deaccelerates the predicted amodal development of words. This suggests an important increasing linear effect and a small quadratic deceleration for the acquisition of the amodal properties of words. Figure 2 presents the predicted developmental trajectories of the amodal representations of words. As it can be seen, the general trend is similar in both languages, except for: (1) the negative quadratic effect is more pronounced for Spanish than English, probably related to some exemplars of English words presenting large quadratic positive trends, and (2) the English data set presented more variability than the Spanish one, probably related to the contents and the sampled ages of the corpuses. In any case, one of the most important findings is that the variability of the predicted amodal development of words was large in both languages, which would be translated into different patterns of semantic change for words: words with small changes across time, words with a mainly linear trend, and words with larger quadratic effects. We hypothesized that part of this variability may be explained by the emotionality of words.

To put into test that hypothesis, we evaluated the main effects of the emotional variables and their interaction with time effects. We found the same pattern of results for the effects of valence and arousal in the trajectories of amodal development of words in

[4] Note that the estimated effects of time were similar in the models with and without emotional norms as covariates.

both languages. In the English data set, the main effect revealed a relevant influence of valence on the initial levels of amodal development (*b*=.406, *SE*=.049, *t*=8.294, *p*<.001), and its interaction effects with both the linear (*b*=-.036, *SE*=.007, *t*=-5.091, *p*<.001) and the quadratic (*b*=.001, *SE*=.000, *t*=5.282, *p*<.001) time effects. On the other hand, the effects of arousal were not relevant (see Table 2). In the Spanish data set, the main effect also revealed a relevant influence of valence on the initial levels of amodal development (*b*=.334, *SE*=.057, *t*=5.841, *p*<.001), and its interaction effects with both the linear (*b*=-.031, *SE*=.008, *t*=-3.822, *p*<.001) and the quadratic (*b*=.001, *SE*=.000, *t*=3.993, *p*<.001) time effects. On the contrary, the effects of arousal were not relevant according to established criteria (see Table 2). These results mean that only emotional valence showed a relevant influence on the amodal development of words. The directionality of this influence is as follows: (1) the more positive emotional words are semantically developed earlier, that is, they present higher initial levels of amodal development (main effect of emotional valence), and (2) the less positive words present more amodal changes through the study, that is, they have more pronounced amodal changes across the study (interaction effects between time and emotional valence), probably because more negative words present less early amodal development and then they have more room for further amodal development.

TABLE 2. Mixed-effects models with crossed random effects (MEMs-CR) for words and computational variables (continuous time metric) using emotional valence and arousal as predictors of the effects of time.

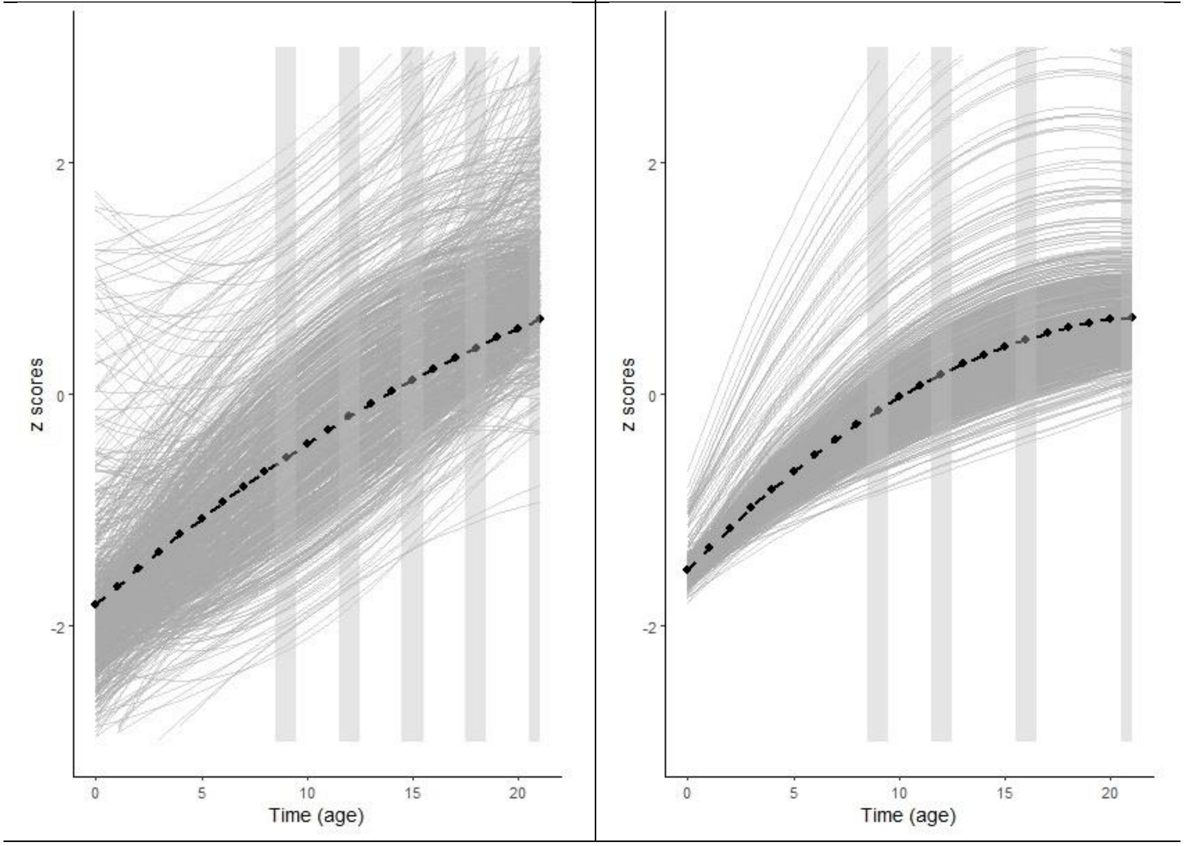
	English words			**Spanish words**		
Fixed effects	**Estimate**	***SE***	***p*-value**	**Estimate**	***SE***	***p*-value**
Intercept	-1.819	.544	.451	-1.519	.568	.008
Time linear effect (TLE)	.159	.007	<.001	.190	.007	<.001
Time quadratic effect (TQE)	-.001	.000	<.001	-.004	.000	<.001
Emotional valence	.406	.049	<.001	.334	.057	<.001
Arousal	-.053	.048	.269	.093	.056	.094
TLE*Emotional valence	-.036	.007	<.001	-.031	.008	<.001
TQE*Emotional valence	.001	.000	<.001	.001	.000	<.001
TLE*Arousal	.003	.007	.713	-.015	.008	.043
TQE*Arousal	.000	.000	.612	.001	.000	.009
Correlation between time effects						
Intercept – TLE	-.083			-.088		
Intercept – TQE	.075			.086		
TLE – TQE	-.984			-.993		
Random effects	**Variance**	***SD***		**Estimate**	***SD***	
Intercepts (words)	.814	.902		.740	.860	
Intercepts (variables)	.879	.938		.961	.980	
Random slopes TLE (words)	.078	.279		.020	.142	
Random slopes TQE (words)	.000	.013		.000	.005	
Model fit	**Deviance**	**AIC**	**BIC**	**Deviance**	**AIC**	**BIC**
	358977	359011	359176	228210	228244	228402

Note. *SE* = Standard error. *SD* = Standard deviation. Given large number of fixed effects, *p*-values were interpreted using α/9, that is, considering *p*<.0056 as significant.

Figure 2

Predicted longitudinal trajectories for the development of amodal representations of English and Spanish words.

A) English words.	B) Spanish words.



Note. Black line: General trajectory. Grey lines: Trajectories of 1000 random words from each data set. x limits: 0 to 21 years. Grey shading: Actual time period covered by the computational study.

DISCUSSION

The aim of the present study was to analyze the influence of emotional valence and arousal in the development of amodal representations of words, combining feature-based and vector space models. In this longitudinal computational study, we evaluated more than 23,000 words, whose emotional norms were available for more than 10,000 words, in each language separately. Results showed a curve of amodal development with an important increasing linear effect and a small quadratic deceleration across early file span. A relevant influence on the amodal development of words was found only for emotional valence, suggesting that emotional words have an earlier amodal development and a less pronounced amodal change during early development. Apart of some quantitative differences, these findings were replicated in both English and Spanish.

Previous research found that young children anchor their emotional processing on valence-based information, and that later development of verbal knowledge increase the multidimensionality of their emotional processing (e.g., Nook et al., 2017). The results of the present study suggest that positive emotional words are semantically developed earlier than negative words, and that their later amodal development is smaller. This might be explained by the early acquisition of positive words and their later smaller change, being the amodal representations of these words almost completely developed relatively soon, at least comparing to negative words. This finding suggests that the semantic networks of positive words would be acquired and maturated earlier than the ones of negative-valence words. For example, it would be expected to find that the semantic networks of words like “family” or “pet” are more mature at earlier developmental stages than “death” or “devil”. In our opinion, there are two different but complementary explanations for these findings:

(1) children could be exposed more frequently to positive words, which would allow them to reiteratively activate the amodal (semantic) representations of those words and to strengthen them; or (2) there might be more complex relationships between language and emotions that could make positive-valence words more likely to be acquired and/or processed during perception, memory or imaginability rather than just activating amodal representations like in abstract words. In fact, recent research has found that earlier-acquired meanings are rated as more concrete than later-acquired meanings, reinforcing the role of sensorimotor experiences in early acquisition of meanings (Muraki et al., in press). Combining both tentative explanations, words to which children are usually exposed could be positive or neutral (more than negative) and then their maturation could be more reinforced thanks to multimodal correlates that link exteroceptive and interoceptive states related to those words. These hypotheses are compatible with previous research on how language might shape emotional experiences (e.g., Hoemann & Feldman-Barrett, 2019) and the later increase of the multidimensionality richness of emotional processing (e.g., Nook et al., 2017). In this sense, positive-valence words would be acquired earlier to anchor a net of meaning construction that would be progressively extended to discriminate between more complex emotional dimensions like high/low arousal and other relevant emotional dimensions.

The main limitation of the present longitudinal computational study is that we constructed the vector spaces using language from linguistic corpora of academic texts, fables, and tales. Future research should try to replicate this study using natural language corpora produced by children. In this sense, although the estimated amodal developmental trajectories were similar to other human cognitive abilities, which might be interpreted as validity evidence of the scores of the computational model, we could expect relevant differences if real and ecologic transcribed language of individuals from different developmental stages. Moreover, it is worth mentioning that the ages of the linguistic corpora were different between both languages. While the statistical models estimate the effects of time in the same continuous time metric, future research should try to use similar developmental stages to make more straightforward comparisons between both languages. Future research might also try to complement this study by adding different emotional properties of words and/or semantic computational measures. Also, we only used the emotional valence and the arousal of words from adults, although we already know that these estimations and the language-emotion relationship might be dynamic -not-static- with complex relationships with amodal properties of words (e.g., Martínez-Huertas, Jorge-Botana, & Olmos, 2021). As the available emotional feature-based data sets for different developmental stages become available in the scientific literature, it will be possible to overcome some of these limitations.

In conclusion, the present longitudinal computational study generated some interesting ideas about the influence of emotions in the amodal development of words. In this line, while most of the previous research asked human raters to evaluate the emotionality of words or concepts at different developmental stages, this study was focused on how some emotional norms influence the initial state and the slope of the amodal development of words or concepts. First, we found that the amodal representations of words follow a developmental trajectory similar to other human cognitive abilities. Second, we found that emotional vocabularies could shape both the initial level of the acquisition of words and influence their change across early life span. Acknowledging that this longitudinal computational model is just an abstraction of actual dynamics of psychological phenomena, we expect that these findings could serve as a thought experiment to discuss about the bidirectional, complex and complementary relationship that connects the emotional processing of words and the acquisition of their amodal meanings.

DATA AVAILABILITY STATEMENT

The amodal (semantic) measures of all the words of the vector space models at each developmental stage and the code for data analysis are available in the OSF project with DOI:10.17605/OSF.IO/9XCT6.

ACKNOWLEDGEMENTS

The authors are thankful to SALSA Lab (Institute of Cognitive Science, University of Colorado, Boulder) for sharing the TASA corpus. We also would like to thank Fritz Günther for his help to obtain the English corpus-based data.

REFERENCES

Fraga, I., Guasch, M., Haro, J., Padrón, I., & Ferré, P. (2018). EmoFinder: The meeting point for Spanish emotional words. *Behavior Research Methods, 50*, 84-93. https://doi.org/10.3758/s13428-017-1006-3.

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science, 14*(6), 1006-1033. https://doi.org/10.1177/1745691619861372.

Hoemann, K., & Feldman-Barrett, L. (2019). Concepts dissolve artificial boundaries in the study of emotion and cognition, uniting body, brain, and mind. *Cognition and Emotion, 33*(1), 67-76. https://doi.org/10.1080/02699931.2018.1535428.

Hoffman, P., Lambon-Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods, 45*(3), 718–730. https://doi.org/10.3758/s13428-012-0278-x

Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review, 125*(3), 293-328. http://dx.doi.org/10.1037/rev0000094.

Ivens, S. H., & Koslin, B. L. (1991). *Demands for Reading Literacy Require New Accountability Methods*. Touchstone Applied Science Associates.

Jorge-Botana, G., Olmos, R., & Barroso, A. (2013, July). *Gallito 2.0: A Natural Language Processing tool to support Research on Discourse*. Proceedings of the Twenty-third Annual Meeting of the Society for Text and Discourse, Valencia.

Jorge-Botana, G., Olmos, R., & Luzón, J. M. (2018). Word maturity indices with latent semantic analysis: why, when, and where is Procrustes rotation applied? *Wiley Interdisciplinary Reviews: Cognitive Science, 9*(e1457), 1-16. https://doi.org/10.1002/wcs.1457.

Jorge-Botana, G., Olmos, R., & Luzón, J. M. (2020). Bridging the theoretical gap between semantic representation models without the pressure of a ranking: some lessons learnt from LSA. *Cognitive Processing, 21*, 1-21. https://doi.org/10.1007/s10339-019-00934-x.

Jorge-Botana, G., Olmos, R., & Sanjosé, V. (2017). Predicting word maturity from frequency and semantic diversity: a computational study. *Discourse Processes, 54*(8), 682-694. https://doi.org/10.1080/0163853X.2016.1155876.

Kireyev, K., & Landauer, T. K. (2011, June). *Word maturity: Computational modeling of word knowledge*. Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies.

Landauer, T. K. & Dumais, S. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–40. https://dx.doi.org/10.1.1.184.4759.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2 3), 259-284. https://doi.org/10.1080/01638539809545028.

Landauer, T. K., Kireyev, K., & Panaccione, C. (2011). Word maturity: A new metric for word knowledge. *Scientific Studies of Reading, 15*(1), 92-108. https://doi.org/10.1080/10888438.2011.536130.

Martínez-Huertas, J. A., & Ferrer, E. (2023). Mixed-effects models with crossed random effects for multivariate longitudinal data. *Structural Equation Modeling, 30*(1), 105-122. https://doi.org/10.1080/10705511.2022.2108430.

Martínez-Huertas, J. A., Jorge-Botana, G., Luzón, J. M., & Olmos, R. (2021). Redundancy, isomorphism and propagative mechanisms between emotional and amodal representations of words: A computational study. *Memory & Cognition, 49*(2), 219–234. https://doi.org/10.3758/s13421-020-01086-6.

Martínez-Huertas, J. A., Jorge-Botana, G., & Olmos, R. (2021). Emotional valence precedes semantic maturation of words: A longitudinal computational study of early verbal emotional anchoring. *Cognitive Science, 45*(e13026), 1-26. https://doi.org/10.1111/cogs.13026.

Martínez-Huertas, J. A., Olmos, R., & Ferrer, E. (2022). Model selection and model averaging for mixed-effects models with crossed random effects for subjects and items. *Multivariate Behavioral Research, 57*(4), 603-619. https://doi.org/10.1080/00273171.2021.1889946.

Muraki, E. J., Reggin, L. D., Feddema, C. Y., & Pexman, P. M. (in press). The Development of Abstract Word Meanings. *Journal of Child Language*. https://doi.org/10.1017/S0305000923000569.

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2017). Increasing verbal knowledge mediates development of multidimensional emotion representations. *Nature Human Behaviour, 1*(12), 881–889. https://doi.org/10.1038/s41562-017-0238-7.

Panksepp, J. (2015). Primal emotions and cultural evolution of language: Primal affects empower words. In U. M. Lüdtke (Ed), E*motion in Language: Theory – research – application* (pp. 27–48). https://doi.org/10.1075/ceb.10.02pan.

Pexman, P. M. (2019). The role of embodiment in conceptual development. *Language, Cognition and Neuroscience*, *34*(10), 1274-1283. https://doi.org/10.1080/23273798.2017.1303522.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178. https://doi.org/10.1037/h0077714.

Sebastián, N., Martí, M. A., Carreiras, M. F., & Cuetos, F. (2000). *LEXESP: Léxico Informatizado del Español [LEXESP: Computerized Lexicon of Spanish].* Edicions Universitat Barcelona.

Shablack, H., & Lindquist, K. A. (2019). The role of language in emotional development. In V. LoBue, K. Pérez-Edgar, & K. A. Buss (Eds), *Handbook of Emotional Development* (pp.451–478). Springer.

Stadthagen-Gonzalez, H., Imbault, C., Pérez-Sánchez, M. A., & Brysbaert, M. (2017). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods, 49*, 111-123. https://doi.org/10.3758/s13428-015-0700-2.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*, 1191-1207. https://doi.org/10.3758/s13428-012-0314-x.