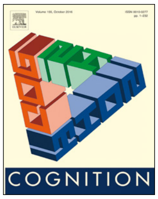
Cognition 190 (2019) 61–71

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit



Original Articles

Sensitivity to emotion information in children's lexical processing

Tatiana C. Lund, David M. Sidhu, Penny M. Pexman*

University of Calgary, Canada

ARTICLE INFO

Keywords:
Emotion
Word valence
Lexical processing
Affective embodiment
Auditory lexical decision
Concreteness

ABSTRACT

We tested predictions of multiple representation accounts of conceptual processing, including the proposal that emotion information may provide a bootstrapping mechanism for vocabulary acquisition. We investigated the influence of word valence on children's lexical processing, presenting 40 positive words, 40 neutral words, and 40 negative words in an auditory lexical decision task (ALDT), along with 120 nonwords. We tested 99 children across three age groups: 5, 6, or 7 years. There were no significant effects of valence on the ALDT responses of 5-year-old children. The 6-year-old children, however, were faster to respond to negative words than to neutral words and, for more abstract words, faster to respond to positive words than to neutral words. The 7-year-old children were faster for positive words than for neutral words, regardless of concreteness. As such, children showed sensitivity to word valence in lexical processing, at a younger age than had been established in previous research. In addition, children's language skills were related to their improved processing of more abstract neutral words between 6 and 7 years of age. These results are consistent with multimodal accounts of word meaning and lexical development.

1. Introduction

According to several recent proposals, conceptual knowledge is acquired and represented in multimodal systems (Barsalou, Santos, Simmons, & Wilson, 2008; Borghi et al., 2017; Dove, 2011, 2018; Thill & Twomey, 2016). That is, word meanings are represented in sensory, motor, emotion, and language systems, and different systems are relatively more important for the representation of different kinds of concepts. These multiple representation views stand in contrast to traditional views which assumed a single system of representation; for instance, that word knowledge is represented in symbolic, amodal format (e.g., Collins & Loftus, 1975) or, alternatively, could be represented in the statistical relationships between words, as captured in lexical co-occurrence (e.g., Lund & Burgess, 1996). The multiple representation views also stand in contrast to strongly embodied accounts, by which it is assumed that all word meanings are grounded in sensorimotor and emotion systems (e.g., Glenberg, 2015; Glenberg & Gallese, 2012).

The multimodal accounts are supported by the results of recent studies with adults, which have shown that responses in simple lexical tasks are influenced by variables that capture the extent to which linguistic, sensory, motor, and/or emotion information is associated with word referents (Moffat, Siakaluk, Sidhu, & Pexman, 2015; Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012). For instance, Yap and Seow (2014; also Kousta, Vinson, & Vigliocco, 2009; Vinson, Ponari, & Vigliocco, 2014) showed that adult participants' responses in a visual lexical decision task (LDT; *is the letter string a real word?*) were affected by word valence, with faster responses for words with positive or negative meanings than for words with neutral meanings. One explanation for facilitatory effects of emotion information is that the emotion information associated with valenced words affords richer semantic representations and thus speeds lexical decisions (Pexman, 2012; Siakaluk et al., 2016).

Some adult studies have shown a somewhat different pattern of valence effects. For instance, in a large-scale analysis of adult LDT responses Kuperman, Estes, Brysbaert, and Warriner (2014) found that responses were fastest to positive words and slowest to negative words, with responses to neutral words falling in between (see also Estes & Adelman, 2008). Although somewhat different to that described above, this pattern still demonstrates sensitivity to emotion information in lexical processing, and has been taken as evidence for automatic vigilance to negative stimuli (Pratto & John, 1991). There is speculation that the particular pattern of valence effects observed in adult studies may depend on stimulus list, small effect sizes, frequency confounds, and other factors that are not yet understood (Kuperman, 2015).

As highlighted in a handful of recent reviews (Marshall, 2016; Pexman, 2018; Wellsby & Pexman, 2014), embodied and multimodal accounts both raise interesting questions about development of word

---

* Corresponding author at: Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N1N4, Canada.
E-mail address: pexman@ucalgary.ca (P.M. Pexman).

https://doi.org/10.1016/j.cognition.2019.04.017
Received 27 June 2018; Received in revised form 10 January 2019; Accepted 17 April 2019

meanings, but their predictions have rarely been tested with children. In the present study, we tested developmental predictions of those accounts by investigating three proposed mechanisms for learning word meanings: (1) the affective embodiment account, (2) the language competence hypothesis, and (3) the nimble-hands, nimble-minds hypothesis.

It has been argued that emotion information might provide an important mechanism for learning and grounding word meaning, particularly for words with less concrete meanings (Barsalou & Wiemer-Hastings, 2005; Glenberg & Gallese, 2012; Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011). Word concreteness is generally defined as the degree to which a word's referent can be experienced through the senses (Brysbaert, Warriner, & Kuperman, 2014), with all words falling somewhere along a continuum from concrete (e.g., *truck*) to abstract (e.g., *truth*). Explaining how we learn and represent the meanings of words that don't have tangible referents has become a central concern in semantic research.

Kousta et al. (2011; also Ponari, Norbury, & Vigliocco, 2017) argued that emotion provides a bootstrapping mechanism for vocabulary acquisition, mapping language to felt experience, because understanding of emotion terms is grounded in bodily experience (Winkielman, Coulson, & Niedenthal, 2018). By this view, valence provides an embodied learning experience for word meanings, especially for abstract word meanings. Borghi et al. (2017) referred to this as the *Affective Embodiment Account*. One way this proposal has been tested is by examining words' age of acquisition (AoA), either from child production norms or from adult ratings of AoA, as an index of concept learning. Results have been mixed, however, with some evidence that words with emotional meanings are acquired earlier (Kousta et al., 2011; Moors et al., 2013) and another study showing that word valence is not related to AoA (Thill & Twomey, 2016).

An alternative strategy for evaluating the Affective Embodiment Account is to directly test whether children's early lexical processing is influenced by word valence. This was the strategy adopted in the present study. To our knowledge, this has been tested in only one previous study. Ponari et al. (2017; also summarized in Vigliocco, Ponari, & Norbury, 2018) conducted an auditory lexical decision task (ALDT) with children aged 6–12 years. A total of 48 word stimuli were selected to achieve a factorial manipulation of valence and concreteness (8 words of each type: concrete positive, concrete neutral, etc.). Children's ALDT reaction times were not measured but analyses of their response accuracy showed that effects of valence were significant only for the 8–9 year old age group, and only for abstract words (effects of valence were marginal for concrete words). Ponari et al. inferred that only children aged 8–9 years were are able to derive significant benefit from valence, but noted that it was difficult to make inferences from the results of the younger children because those children had very low accuracy in the ALDT (< 70%). The word stimuli selected by Ponari et al. had quite high mean AoA (in years; $M = 7.80$), suggesting that the younger children may not have known as many of the words' meanings as did the older children. As such, Ponari et al. concluded that floor effects might have been an issue in this younger age group. Thus, to evaluate the Affective Embodiment Account, there is a need to test younger children with age-appropriate words to determine whether they are sensitive to valence information in lexical processing.

Another explanation for children's word learning is that they are able to derive meanings from their experience with language, including their knowledge of co-occurrence information in linguistic input (Vigliocco et al., 2018). That is, aspects of a word's meaning can be extracted from the linguistic context in which it is used. Vigliocco et al. (2018) argued that while this information would be relevant for learning both concrete and abstract words, it would be a particularly important mechanism for learning abstract words, for which meanings cannot be mapped to referents in the physical environment (also Vigliocco, Meteyard, Andrews, & Kousta, 2009; Vigliocco, Ponari, & Norbury, 2017). Similarly, it has been argued that the meanings of neutral words, particularly neutral words without concrete referents, could be acquired through language, by their use in the context of other words (Howell, Jankowicz, & Becker, 2005).

As such, linguistic experience should be important to children's acquisition of abstract word meanings, and we refer to this here as the *language competence hypothesis*. To our knowledge, this possibility has been tested in only one previous study. Ponari, Norbury, Rotaru, Lenci, and Vigliocco (2018; also reported in Vigliocco et al., 2017) examined whether children with developmental language disorder (DLD) found abstract words to be particularly challenging. Despite having the expected language deficits, the children with DLD did not show disproportionately poor performance for abstract words relative to concrete words. The authors took these findings as evidence that language competence is not the primary driver of abstract word learning. Instead, they inferred that both typically developing children and children with DLD are likely able to learn abstract and concrete meanings through affective and sensorimotor associations.

Vigliocco et al. (2018) argued, further, that the Affective Embodiment Account could be complementary with a role for language competence, for development of an abstract vocabulary. In particular, "A likely scenario is one in which while emotional grounding plays a key role early in development, linguistic information becomes essential later on" (p. 537). The suggestion is that emotion information might help children to begin to represent the differences between abstract and concrete words. In particular, the fact that abstract emotion terms refer to internal states may help children to create a framework for abstract concepts (Ponari et al., 2018). Once this framework is established, language competence becomes the key factor to help children derive word meanings from patterns in language use (Vigliocco et al., 2017). As yet, there is no empirical evidence for this proposed tradeoff in abstract word learning, wherein there is early reliance on emotion information and later support from language skills, but we tested this proposal in the present study, by investigating the relationship of children's language skills to their reliance on emotion information in lexical processing of abstract words.

In addition to linguistic and emotion information, multiple representation views of conceptual knowledge propose that sensorimotor information is important to understanding word meanings (e.g., Barsalou et al., 2008; Dove, 2018). In support of this position, there is evidence that children's early fine motor or manual dexterity skills are related to their language skills (e.g., Grissmer, Grimm, Aiyer, Murrah, & Steele, 2010; Pexman & Wellsby, 2016). Two recent studies have provided more specific support for the role of sensorimotor information in vocabulary acquisition. Suggate and Stoeger (2014) tested the *nimble-hands, nimble-minds hypothesis*: the notion that children who have more advanced fine motor skills will have richer sensorimotor interactions with objects, and thus will show more advanced understanding for terms referring to objects that afford easy interaction. To quantify this aspect of object information, Suggate and Stoeger used the body-object-interaction (BOI) dimension (Siakaluk, Pexman, Aguilera, Owen, & Sears, 2008). High BOI words refer to objects with which the human body can easily interact, whereas low BOI words refer to objects with which the human body can less easily interact. In a sample of 3–6 year old children, Suggate and Stoeger found that children's fine motor skills were more strongly related to their accuracy on a receptive vocabulary test for high BOI words than for low BOI words. In a subsequent study, Suggate and Stoeger (2017) also found that children's fine motor skills were related to their speed to point to pictures of high BOI objects, in that children with more advanced fine motor skills tended to point to pictures of named high BOI objects more quickly than did children with less advanced fine motor skills. They concluded that children's vocabulary development is initially grounded in their sensorimotor experience before it becomes more abstract. While these relationships between fine motor skills and knowledge of high BOI words have been observed by Suggate and Stoeger in receptive vocabulary accuracy and pointing latencies, they have not been tested in the conventional lexical processing task (lexical decision, or auditory lexical decision as in the Ponari et al., 2017, study) but we did so here.

In the present study we examined children's lexical processing, measuring their responses in an ALDT. We focused on younger children than in the Ponari et al. (2017) study (groups of 5-, 6-, and 7-year-olds), using a large set of words (40 positive words, 40 neutral words, 40 negative words) that were chosen to be familiar for children in this age range (AoA $M = 5.35$). With the larger word set, we hoped to have enough correct ALDT responses to permit analyses of both reaction times (the primary behavioral measure in the lexical processing literature) and accuracy. The words within each valence type were matched for frequency and several other lexical dimensions and varied in concreteness values. We did not manipulate concreteness categorically, in the way that Ponari et al. did, because we found this wasn't feasible given the younger age range we wanted to test, the large number of items we wanted to include, and the lexical factors we wanted to match across word types. It is estimated that the average 5-year-old vocabulary contains less than 20% abstract words, and that children of this age are just beginning to develop their knowledge of abstract word meanings (Ponari et al., 2017). This limits the number of familiar words with very abstract meanings to choose from. In addition, we wanted to match our stimuli for AoA, yet AoA tends to be higher for abstract words than for concrete words (Ponari et al., 2017; Thill & Twomey, 2016), and our preliminary checks on the potential item set suggested that the matching of AoA for concrete and abstract words would not be possible. As such, we elected to examine the effects of valence on children's lexical processing across positive, neutral, and negative word types, and included both concreteness and AoA as continuous variables in our analyses. We did not include frequency and other lexical variables as predictors in the analyses because these variables were not significantly different for more concrete and more abstract words (all $p > .27$). Given the age-sensitive nature of the valence effects observed in Ponari et al. (recall that they found valence effects in only one intermediate age group), we expected that valence effects might be somewhat different in the three age groups tested. To fully evaluate this possibility, we planned analyses of each age group (5-, 6-, and 7-year-olds) separately for effects of word valence. We also tested the interaction of age and valence in the overall analyses. Further, based on the Affective Embodiment Account, we hypothesized that there might be an interaction of valence and concreteness. Indeed, Ponari et al. found somewhat different valence effects for concrete and abstract words. Finally, we assessed children's language and fine motor skills, and these measures allowed us to test additional theoretical predictions, outlined above, related to the language competence hypothesis and the nimble-hands, nimble-minds hypothesis.

2. Method

2.1. Participants

A total of 99 children participated in the study, including 35 5-year-old children ($M = 5;6$, $SD = 0;2$, 21 female), 34 6-year-old children ($M = 6;5$, $SD = 0;3$, 15 female), and 30 7-year-old children ($M = 7;5$, $SD = 0;3$, 15 female). All children were recruited through our participant database and received a small toy and a pencil case for participating. Two participants (one 5-year-old female and one 6-year-old female) were excluded from the analyses for failing to score above chance for ALDT accuracy (chance performance was 50% accuracy). Ten additional participants were excluded by a response criterion described in the Results section. Thus, 87 participants were included in the analyses.

2.2. Stimuli

Word stimuli for the ALDT were 40 positive words (e.g., *cake*, *heart*), 40 neutral words (e.g., *map*, *nest*), and 40 negative words (e.g., *jail*, *trash*). Word valence was determined based on published norms (Warriner, Kuperman, & Brysbaert, 2013). All words were monosyllabic. Words were selected such that the three word sets differed in mean valence, but were not significantly different on a number of other factors that might influence ALDT performance (Table 1): number of phonemes, phonological Levenshtein distance (PLD, a measure of words' phonological similarity or confusability, Yarkoni, Balota, & Yap, 2008), children's spoken frequency for the 72–83 month age range (from the ChildFreq norms, which are extracted from the CHILDES database, as described in Bååth, 2010), Grade 2 print frequency (Zeno, Ivens, Millard, & Duvvuri, 1995), adult log SUBTLEX$_{WF}$ word frequency (Brysbaert & New, 2009), age of acquisition (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012), imageability (Cortese & Fugett, 2004), and concreteness (Brysbaert et al., 2014).

Table 1
Mean characteristics of word stimuli, as a function of word type (standard deviations in parentheses).

	Word type			
	Positive	Neutral	Negative	p for effect of word type
Word characteristics				
Valence	7.18 (0.43)	5.47 (0.87)	3.14 (0.68)	< .001
Number of phonemes	3.33 (0.62)	3.50 (0.68)	3.35 (0.62)	.42
PLD	1.26 (0.28)	1.29 (0.32)	1.22 (0.27)	.56
Child spoken frequency	76.25 (104.97)	108.63 (171.24)	71.98 (90.73)	.38
Grade 2 print frequency	85.55 (60.44)	87.73 (72.61)	81.18 (61.37)	.90
Adult word frequency	3.58 (0.52)	3.45 (0.45)	3.55 (0.50)	.49
Age of acquisition	5.16 (1.18)	5.62 (1.57)	5.26 (1.23)	.27
Imageability	4.64 (1.48)	4.67 (1.45)	4.64 (1.21)	.99
Concreteness	3.45 (1.20)	3.72 (1.05)	3.78 (0.86)	.34

Note. PLD = phonological Levenshtein distance.

In addition, we selected a set of 120 monosyllabic nonwords (e.g., *darb*). We chose nonwords from the English Lexicon Project (Balota et al., 2007) such that they were matched to the word stimuli for onset phoneme and number of phonemes, so that these factors could not be used as a cue to the decision.

A female speaker recorded word and nonword ALDT stimuli in a sound attenuated chamber. Once recorded, sound files were edited with the program Praat (Boersma, 2001) to ensure that there were no significant differences in sound file length across word types, or between words and nonwords. Three adult listeners who were naïve to the study purpose listened to the sound files and reported what they heard to ensure that each file was interpreted as intended.

2.3. Procedure

Participants were tested in our university laboratory. They sat in front of a computer wearing headphones. The experimenter sat beside the participant, and also wore headphones. Sound files were presented one at a time through both the participant's and the experimenter's headphones using E-Prime presentation software (Schneider, Eschman, & Zuccolotto, 2001). Sound files were presented in a different random order for each participant. The computer screen was blank except for a small central fixation cross. Participants were told that they were playing a game in which they were word detectives. A response box with red and green response buttons was placed in front of participants. Participants were instructed to place their index fingers on the red and green buttons, and to press the red button when they heard a fake word and the green button when they heard a real word, pressing the button as soon as they decided. To ensure understanding of the game, participants were first presented with a practice block consisting of ten trials (five words and five nonwords). After each response, the experimenter pressed a button on a standard keyboard to proceed to the next item.

Following the ALDT, children completed the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007) and manual dexterity subscales of the Bruininks-Oseretsky Test of Motor Profiency-2nd Edition (BOT-2; Bruininks & Bruininks, 2005). The BOT includes five manual dexterity subscales: drawing dots in circles (BOT-Dots), picking up pennies and transferring them into a box (BOT-Pennies), placing small pegs into a board (BOT-Pegs), sorting cards (BOT-Cards), and stringing blocks on a shoelace (BOT-Blocks). Each subscale task is performed in two 15-second trials (excluding dots which is performed once). Parents completed the Children's Communication Checklist-2 (CCC-2; Bishop, 2006). This caregiver-report measure is comprised of 70 items divided into 10 scales (A: Speech, B: Syntax, C: Semantics, D: Coherence, E: Initiation, F: Scripted language, G: Context, H: Nonverbal communication, I: Social relations, J: Interests) to assess children's communication difficulties and pragmatic language. Caregivers are asked to rate the frequency with which their child displays each item's statement, ranging from 0 (less than once a week, or never) to 3 (several times i.e. more than twice a day, or always). A General Communication Composite (GCC) score is calculated by summing scaled scores of scales A-H. Four participants' CCC-2 Checklists (data for two 5-year-old and two 6-year-old participants) were not scored due to incomplete or missing parent responses. Mean scores for each task are included in Table 2.

2.4. Supplementary material

The data analyzed in this study are available here: https://osf.io/9a5ng/.

3. Results

Analyses consisted of mixed effects regression models. Models were computed using the "lme4" package (Bates, Maechler, Bolker, & Walker, 2015) in the statistical software R (R Core Team, 2017). In each model, we took a confirmatory approach, and fit all fixed effects of interest. We developed each model's random effects structure using the approach suggested by Bates, Kliegl, Vasishth, and Baayen (2015). This was as follows:

1. We began with a model containing all possible random effects. For cases in which this did not converge, we fit a simpler model that omitted correlations among random effects. This was done using the "afex" package in R (Singmann, Bolker, & Westfall, 2015).
2. We then used the "RePsychLing" package in R (Baayen, Bates, Kliegle, & Vasishth, 2015) to perform a principal components analysis on this random effects structure to determine the number of random effects that could be specified (i.e., the number of components explaining > 1% of variance) while achieving model identification. Beginning with the highest order random effect with the least amount of variance, we removed random slope effects until we reached a model that contained the number suggested by the principal components analysis.
3. If correlations among random effects were retained up to this point, we then compared models with and without correlations among random effects using likelihood ratio tests (*LRTs*) to determine if they were warranted.
4. We then tested the inclusion of every remaining random slope effect, beginning with the highest order effect with the least amount of variance, using LRTs.
5. Finally, if there were no correlations among random effects, we tested whether the model could be improved by their inclusion using LRTs.

Code for this entire process, for each model, is available here: https://osf.io/9a5ng/. We only report the results of the model containing the final random effects structure in the text. Note that models always included random subject and item intercepts to deal with non-independence. Continuous predictors were always mean-centered.

Analyses were only carried out on real word trials as these involve the manipulation of word valence. Nonword response data were not analyzed further, but mean latencies and accuracy for nonword responses are presented in Table 2. We removed three words with an overall accuracy below 50%: *rise* (49.48%), *tramp* (42.27%), and *whip* (38.14%), from all analyses. In addition, as in Ponari et al. (2017), we conducted a signal detection analysis to determine whether any of the participants had a bias towards responding either "word" or "nonword". We computed a Criterion C for each participant, using "word" responses to real words as hits, and "word" responses to nonwords as false alarms. We excluded participants with a Criterion C value greater or less than 1.5 standard deviations from the mean of their age group. This led to the exclusion of four 5-year-olds, four 6-year-olds, and two 7-year-olds. Lastly, we removed all trials with a reaction time less than 200 ms (0.12% of remaining trials) or greater than 3000 ms (5.46% of remaining trials) as we judged these to be outliers.

3.1. Reaction times

We used mixed effects linear regression models to analyze children's reaction times. These analyses were only carried out on trials that resulted in a correct response. In addition, after removing incorrect trials, we removed trials with latencies greater than 2.5 standard deviations from each participant's mean (2.73% of remaining trials). We used the "lmerTest" package (Kuznetsova, Brockhoff, & Christensen, 2017) to generate *p*-values for models' fixed effects.

3.1.1. Affective Embodiment Account

Our first set of analyses explored the relationship between valence and concreteness in reaction times.

3.1.1.1. All ages. We began with an analysis of reaction times for all age groups. See Table 2 for average reaction times by age group and

Table 2
Mean participant characteristics and ALDT responses (standard deviations in parentheses), as a function of age group (N = 87).

	5-year-olds (n = 30)	6-year-olds (n = 29)	7-year-olds (n = 28)
PPVT4 (Raw Score)	117.60 (15.53)	132.59 (15.77)	145.11 (16.13)
BOT2 (Point Score)	16.47 (3.08)	19.76 (2.50)	22.46 (3.65)
CCC2 GCC	82.36 (11.31)	81.70 (15.07)	85.79 (9.99)
ALDT word accuracy (%)	78.47 (14.65)	90.33 (5.93)	94.32 (3.61)
ALDT word reaction time (ms)	1473.33 (235.31)	1364.88 (133.47)	1329.40 (181.29)
ALDT nonword accuracy (%)	72.45 (18.80)	87.25 (11.70)	86.97 (10.99)
ALDT nonword reaction time (ms)	1605.23 (306.62)	1618.63 (205.80)	1601.95 (201.78)

Note. PPVT4 = Peabody Picture Vocabulary Test-4th Edition; BOT2 = Bruininks-Oseretsky Test of Motor Proficiency-2nd Edition; CCC2 GCC = Children's Communication Checklist-2 General Communication Composite; ALDT = Auditory Lexical Decision Task. Means and standard deviations calculated for trials and participants included in the analyses. Note that nonword trials were cleaned in the same manner described below for real word trials.

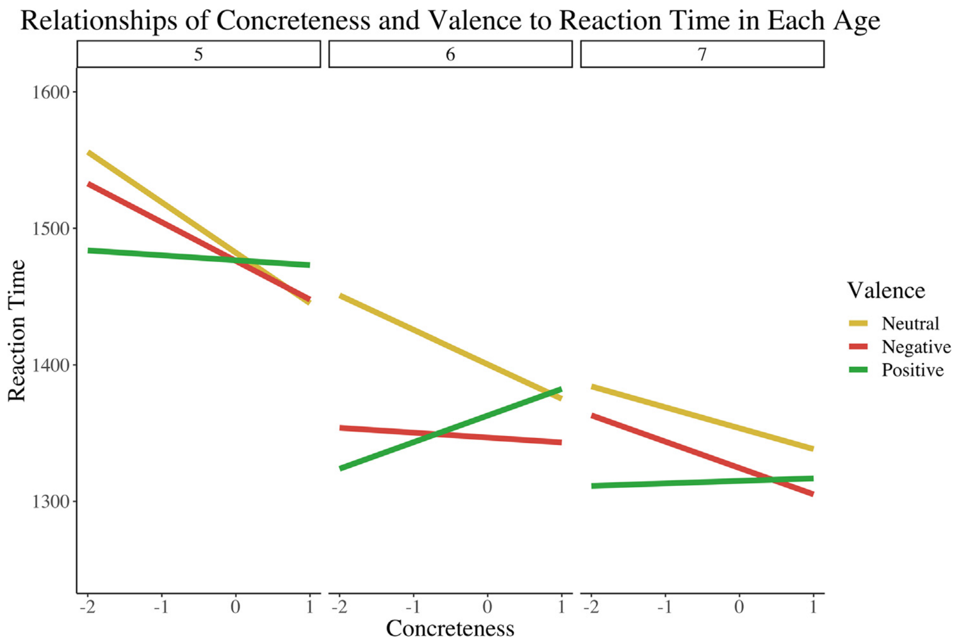


Fig. 1. Marginal plots for relationships of standardized concreteness and valence to reaction time in each age group.

Table 3
Linear mixed effects regression model predicting reaction time for all ages, including interactions.

Fixed Effect	*B*	*S.E.*	*t*	*p*
Intercept	1354.10	45.22	29.94	< .001***
Control Variables				
Age of acquisition	13.52	6.41	2.11	.04*
Voicing	−13.31	16.95	−0.79	.43
Bilabial	67.30	35.49	1.90	.06
Labiodental	97.63	31.21	3.13	.002**
Alveolar	87.03	30.56	2.85	.005**
Palatal	109.75	42.08	2.61	.01*
Velar	77.09	36.85	2.09	.04*
Stop	−31.63	23.13	−1.37	.17
Fricative	14.70	32.39	0.45	.65
Affricate	−100.17	30.00	−3.34	.001**
Nasal	12.29	25.01	0.49	.62
Uniqueness Point	10.83	6.28	1.73	.09
Predictor Variables				
Concreteness	−25.34	10.94	−2.32	.02*
Age (6)	−80.74	50.95	−1.59	.12
Age (7)	−45.80	50.42	−0.91	.37
Valence (Negative)	−29.64	14.18	−2.09	.04*
Valence (Positive)	−27.19	14.03	−1.94	.06
Concreteness × Valence (Negative)	8.75	16.71	0.52	.60
Concreteness × Valence (Positive)	31.57	13.52	2.34	.02*
Age (6) × Valence (Negative)	−46.18	26.89	−1.72	.09
Age (7) × Valence (Negative)	22.04	21.29	1.04	.30
Age (6) × Valence (Positive)	−37.53	26.56	−1.41	.16
Age (7) × Valence (Positive)	1.59	21.27	0.08	.94
Random Effect	s^2			
Item Intercept	2055.33			
Item Age (6) Slope	4249.20			
Item Age (7) Slope	0			
Subject Intercept	32923.83			
Residual	107213.22			

Note. Observations = 8200; Items = 117; Subjects = 87.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Fig. 1 for overall patterns. Age of acquisition (Kuperman et al., 2012) and uniqueness point (Luce, 1986) were included as control variables. We also coded the onset phoneme of each word, following Balota, Cortese, Sergent-Marshall, Spieler, and Yap (2004): we dummy coded variables for voicing, place of articulation (i.e., one each coding for whether the onset phoneme was a bilabial, labiodental, alveolar, palatal, velar or glottal) and manner of articulation (i.e., one each coding for whether the phoneme was a stop, fricative, affricate, nasal or liquid/glide). Note that glottal and liquid/glide terms were excluded automatically due to rank deficiency. Our variables of interest were: valence (neutral, negative or positive; dummy coded using neutral words as a reference category), concreteness (Brysbaert et al., 2014), and age (5-, 6- or 7-year-olds; successive difference coded comparing successive ages). Our a priori hypotheses led us to also include an interaction between concreteness and valence, and between age and valence. This analysis revealed a significant interaction between concreteness and positive (vs. neutral) valence ($p = .02$). No other interactions reached statistical significance (all $p > .09$), see Table 3. Plotting this interaction suggests that positive valence is facilitatory for more abstract words, but not for more concrete words (see Fig. 2). This was confirmed by follow up analyses which split items based on median concreteness. We built a model in each group of items including all previously mentioned control variables, as well as age, with the predictor of interest being valence. These analyses revealed a significant effect of positive (vs. negative) valence for more abstract words ($p = .02$), but not for more concrete words ($p = .62$)

Next, the interaction terms were removed to allow interpretation of individual valence, concreteness, and age predictors. This revealed a significant effect of age (6 vs. 5) in which 6-year-olds responded faster than 5-year-olds ($p = .03$). There was also a significant effect of negative (vs. neutral) valence, in which responses were faster to negative words than to neutral words ($p = .045$). While there was also a significant effect of positive (vs. neutral) valence, this will not be interpreted as it was previously shown to interact with concreteness. See Table 4.

3.1.1.2. *5-year-olds*.

Planned analyses for each age group included all previously mentioned control variables. Predictors of interest were valence, concreteness, and their interaction. For 5-year-olds, no interactions reached statistical significance (all $p > .14$). Next, the

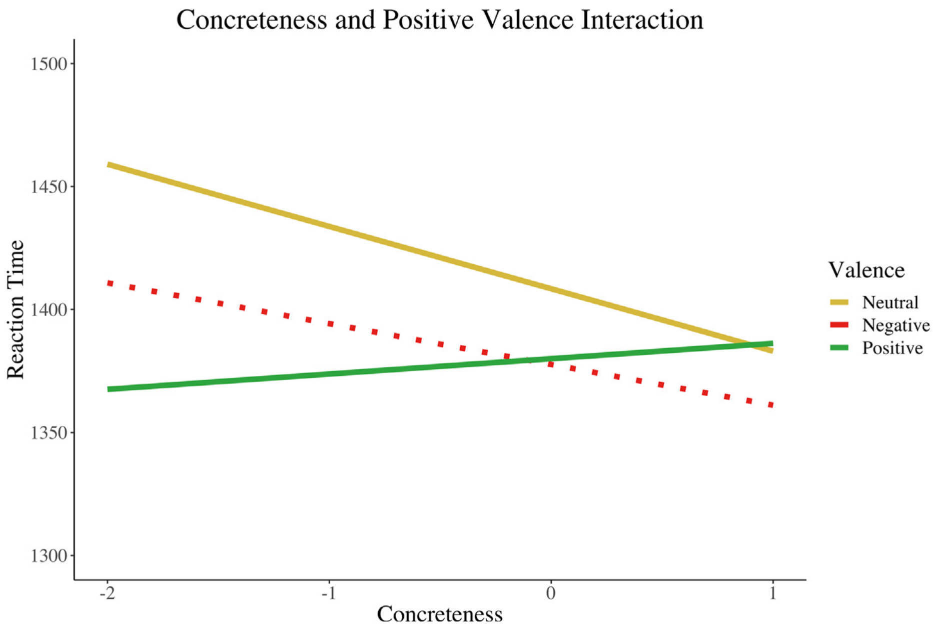


Fig. 2. Marginal plot for interaction between standardized concreteness and positive (vs. neutral) valence in the prediction of reaction time. Negative (vs. neutral) valence also shown for reference, but in dashed form as not part of significant interaction.

Table 4
Linear mixed effects regression model predicting reaction time for all ages, without interactions.

Fixed Effect	*B*	*S.E.*	*t*	*p*
Intercept	1346.61	45.72	29.45	< .001***
Control Variables				
Age of acquisition	14.66	6.23	2.35	.02*
Voicing	−14.90	17.32	−0.86	.39
Bilabial	73.94	36.25	2.04	.04*
Labiodental	102.55	31.88	3.22	.002**
Alveolar	91.99	31.22	2.95	.004**
Palatal	123.11	42.69	2.88	.005**
Velar	81.92	37.63	2.18	.03*
Stop	−28.86	23.45	−1.23	.22
Fricative	14.94	32.84	0.46	.65
Affricate	−105.18	30.44	−3.46	.001**
Nasal	11.64	25.57	0.46	.65
Uniqueness Point	12.47	6.38	1.95	.05
Predictor Variables				
Concreteness	−8.88	6.35	−1.40	.17
Age (6)	−108.90	48.51	−2.25	.03*
Age (7)	−37.81	48.83	−0.77	.44
Valence (Negative)	−29.08	14.34	−2.03	.045*
Valence (Positive)	−28.72	14.28	−2.01	.047*
Random Effect	s^2			
Item Intercept	2228.64			
Item Age (6) Slope	4529.14			
Item Age (7) Slope	0			
Subject Intercept	32896.64			
Residual	107240.84			

Note. Observations = 8200; Items = 117; Subjects = 87.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

interaction term was removed to allow interpretation of individual valence and concreteness predictors. None of these predictors reached statistical significance (all $p > .07$).

3.1.1.3. 6-year-olds. The same analysis was conducted on reaction times for children in the 6-year-old group. This analysis revealed an interaction between concreteness and positive (vs. neutral) valence ($p = .02$). Plotting this interaction suggests that positive valence is facilitatory for more abstract words, but not for more concrete words (see Fig. 3). This was confirmed by follow up analyses which split items based on median concreteness. We built a model in each group of items including all previously mentioned control variables, with the predictor of interest being valence. These analyses revealed a significant effect of positive (vs. neutral) valence for more abstract words ($p = .01$), but not for more concrete words ($p = .57$). The interaction between concreteness and negative (vs. neutral) valence did not reach statistical significance ($p = .39$). In addition, this analysis revealed a simple effect (i.e., at mean levels of concreteness) of negative (vs. neutral) valence, in which children responded faster to negative words than to neutral words ($p = .01$). No other predictors reached statistical significance (all $p > .08$), see Table 5.

3.1.1.4. 7-year-olds. The same analysis was conducted on reaction times for children in the 7-year-old group. No interactions in this model reached statistical significance (all $p > .39$). Next, the interaction terms were removed to allow interpretation of individual valence and concreteness predictors. This revealed a significant effect of positive (vs. neutral) valence in which children responded faster to positive words than neutral words ($p = .02$). No other predictors reached statistical significance (all $p > .10$), see Table 6.

3.1.2. Language competence hypothesis

We next conducted analyses including only the more abstract of our items, based on a median split of concreteness ratings. These analyses included data for children in the 6- and 7-year-old groups, as there was no evidence of valence playing a role for 5-year-olds. We examined whether language competence contributed to differences in the processing of abstract items of different valences. As a first step, to identify a dimension of language competence, we performed a principal components analysis of scores on the PPVT and each of the component subscales of the CCC. We used an oblimin rotation as we expected components to be correlated with one another. We extracted components until a component had an Eigenvalue lower than 1.00. This resulted in three components being extracted (see Table 7 for the pattern matrix).

We used component 3 to quantify language competence, as it included the child's syntactic skills, their capacity to produce fluent and

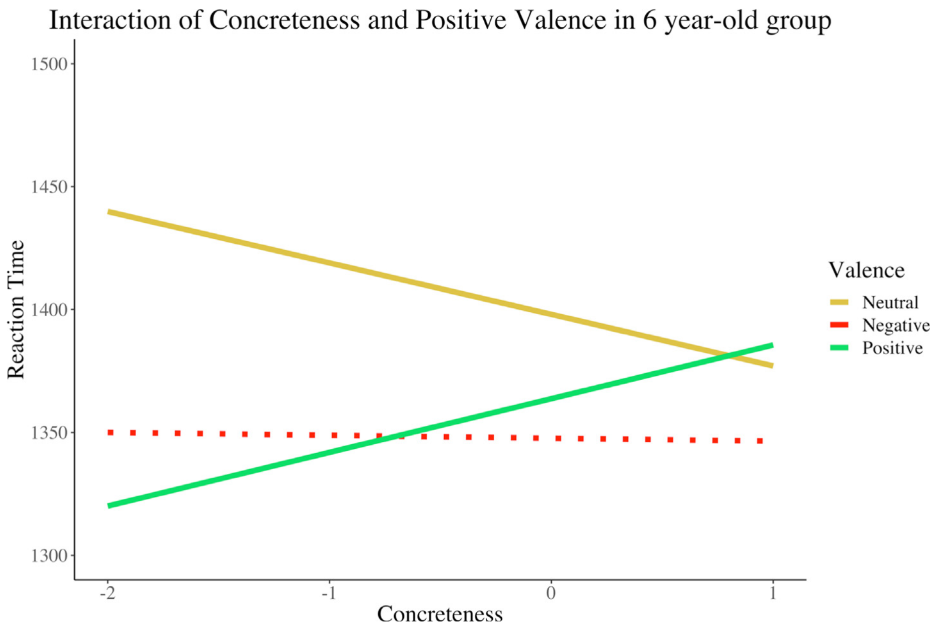


Fig. 3. Marginal plot for interaction between standardized concreteness and positive (vs. neutral) valence in the prediction of reaction time for 6-year-olds. Negative (vs. neutral) valence also shown for reference, but in dashed form as not part of significant interaction.

Table 5
Linear mixed effects regression model predicting reaction time for 6-year-olds.

Fixed Effect	*B*	*S.E.*	*t*	*p*
Intercept	1382.22	60.62	22.80	< .001***
Control Variables				
Age of acquisition	20.52	8.86	2.32	.02*
Voicing	1.85	23.42	0.08	.94
Bilabial	18.25	48.40	0.38	.71
Labiodental	47.85	42.58	1.12	.26
Alveolar	40.99	41.55	0.99	.33
Palatal	97.31	57.72	1.69	.10
Velar	17.15	50.34	0.34	.73
Stop	−28.22	31.85	−0.89	.38
Fricative	−2.30	44.54	−0.05	.96
Affricate	−132.79	41.33	−3.21	.002**
Nasal	−11.09	34.42	−0.32	.75
Uniqueness Point	4.73	8.61	0.55	.58
Predictor Variables				
Concreteness	−20.93	15.16	−1.38	.17
Valence (Negative)	−50.36	19.50	−2.58	.01*
Valence (Positive)	−34.26	19.36	−1.77	.08
Concreteness × Valence (Negative)	19.75	23.03	0.86	.39
Concreteness × Valence (Positive)	42.78	18.70	2.29	.02*
Random Effect	s^2			
Item Intercept	3056.10			
Subject Intercept	16091.36			
Residual	90922.77			

Note. Observations = 2857; Items = 117; Subjects = 29.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 6
Linear mixed effects regression model predicting reaction time for 7-year-olds, without interactions.

Fixed Effect	*B*	*S.E.*	*t*	*p*
Intercept	1291.17	59.13	21.84	< .001***
Control Variables				
Age of acquisition	13.60	7.38	1.84	.07
Voicing	–23.76	20.56	−1.16	.25
Bilabial	56.32	42.80	1.32	.19
Labiodental	89.36	37.72	2.37	.02*
Alveolar	87.35	36.88	2.37	.02*
Palatal	121.03	50.34	2.40	.02*
Velar	70.50	44.51	1.58	.12
Stop	−16.78	27.64	−0.61	.55
Fricative	18.71	38.73	0.48	.63
Affricate	−69.21	36.08	−1.92	.06
Nasal	20.83	30.29	0.69	.49
Uniqueness Point	9.09	7.58	1.20	.23
Predictor Variables				
Concreteness	−10.49	7.52	−1.40	.17
Valence (Negative)	−27.87	16.90	−1.65	.10
Valence (Positive)	−41.51	16.93	−2.45	.02*
Random Effect	s^2			
Item Intercept	2019.78			
Subject Intercept	30865.36			

Note. $^{**}p < .01$. Observations = 2881; Items = 117; Subjects = 28.
* $p < .05$.
*** $p < .001$.

coherent expressions, and to engage in meaningful conversation (Bishop, 1998).[1] The analysis also included all previously mentioned control variables, as well as Age (effects coded, age 6 = −1, 7 = 1) which served as a control variable in the present analysis. Predictors of interest were component 3 scores (henceforth *language competence*), valence, and their interaction. Note that two children who did not have subscale scores on the CCC were not included in this analysis. This analysis revealed a significant interaction between language competence and negative (vs. neutral) valence (*p* = .006). There was not a significant interaction between language competence and positive (vs. neutral) valence (*p* = .70), see Table 8. Plotting this interaction suggests that language competence is facilitatory for neutral abstract words but not for negatively valenced abstract words (see Fig. 4). We used the "jtools" package (Long, 2018) in R to conduct simple slope analyses for language competence, in neutral and negatively valenced words. This revealed a significant facilitatory effect of language competence for neutral items (*p* < .001), and a significant inhibitory effect of language competence for negative items (*p* < .001).

---

[1] Supplementary analyses using the first and second components found that they did not interact with valence.

Table 7
Resulting pattern matrix of the PCA.

Variable	Component 1	Component 2	Component 3
PPVT		0.59	
CCC: Speech			−0.96
CCC: Syntax			−0.84
CCC: Semantics			
CCC: Coherence			−0.55
CCC: Initiation	0.52		
CCC: Scripted Language			
CCC: Context	0.52		
CCC: Nonverbal Communication	0.71		
CCC: Social Relations	0.94		
CCC: Interests		−0.86	

Note: Only loadings > 0.5 are shown.

Table 8
Linear mixed effects regression model predicting reaction time for abstract items, for 6- and 7-year-olds.

Fixed Effect	*B*	*S.E.*	*t*	*p*
Intercept	1319.61	62.97	20.96	< .001***
Control Variables				
Age of acquisition	18.23	9.10	2.00	.049*
Voicing	1.33	27.57	0.05	.96
Bilabial	29.77	50.93	0.59	.56
Labiodental	56.85	44.13	1.29	.20
Alveolar	53.75	44.43	1.21	.23
Palatal	126.01	56.32	2.24	.03*
Velar	112.78	56.59	1.99	.05*
Stop	−8.98	31.49	−0.29	.78
Fricative	54.15	45.68	1.19	.24
Affricate	−81.52	41.89	−1.95	.06
Nasal	6.04	34.54	0.18	.86
Uniqueness Point	−10.79	9.37	−1.15	.25
Age	−13.42	21.59	−0.62	.54
Predictor Variables				
Language Competence	−35.01	22.71	−1.54	.13
Valence (Negative)	−34.43	21.80	−1.58	.12
Valence (Positive)	−54.72	20.71	−2.64	.01*
Language Competence × Valence (Negative)	37.47	13.67	2.74	.006***
Language Competence × Valence (Positive)	5.16	13.28	0.39	.70
Random Effect	s^2			
Item Intercept	1977			
Subject Intercept	23,611			

Note. $^{**}p < .01$. Observations = 2813; Item = 60; Subject = 55.
* $p < .05$.
*** $p < .001$.

3.1.3. *Nimble-hands nimble-minds*

We next conducted an analysis that included all items for which body-object interaction ratings (*BOI*; Pexman, Muraki, Sidhu, Siakaluk, & Yap, 2019) were available (104 words). The model included all previously mentioned control variables as well as Age. Predictors of interest were BOI rating, BOT score, and their interaction. The analysis revealed that the interaction was not significant ($p = .80$). The main effects of BOI ($p = .58$) and BOT ($p = .19$) were also not significant.[2]

3.2. *Response accuracy*

We used mixed effects logistic regression models to analyze children's response accuracy. We removed trials with latency greater than 2.5 SD from each participant's mean (2.65% of remaining trials). In the 7-year-old age group, 14 of the 28 participants had mean accuracy greater than 95% (compared to two 5-year-olds and five 6-year-olds). Given this ceiling effect in accuracy for the 7-year-olds we did not include their accuracy data in the analyses.

3.2.1. *Affective Embodiment*

Our first set of accuracy analyses explored the relationship between valence and concreteness in response accuracy.

3.2.1.1. *Both ages*. We began with an analysis of accuracy data for 5- and 6 year-old children. See Table 2 for average response accuracy by age group and Fig. 5 for overall patterns. The model included all previously mentioned control variables. Our variables of interest were: valence, concreteness (Brysbaert et al., 2014), and age (effects coded, age 5 = −1, 6 = 1). Our a priori hypotheses led us to also include an interaction between concreteness and valence, and between age and valence. This analysis revealed a marginally significant interaction between age and negative (vs. neutral) valence ($p = .07$). No other interactions reached statistical significance (all $p > .46$). Next, the interaction terms were removed to allow interpretation of individual valence, concreteness and age predictors. This revealed a significant effect of age, in which 6-year-olds were more accurate than 5-year-olds ($p < .001$). In addition, there was a marginal effect of negative (vs. neutral) valence, in which children were marginally more accurate in their responses to negative than neutral words ($p = .06$).

3.2.1.2. *5-year-olds*. Planned analyses for each age group included all previously mentioned control variables. Predictors of interest were valence, concreteness, and their interaction. No interactions reached statistical significance (all $p > .53$). Next, the interaction term was removed to allow interpretation of individual valence and concreteness predictors. None of these predictors reached statistical significance (all $p > .18$).

3.2.1.3. *6-year-olds*. The same analysis was conducted on accuracy for the 6-year-old group. No interactions reached statistical significance (all $p > .16$). Next, the interaction term was removed to allow interpretation of individual valence and concreteness predictors. This revealed a marginally significant effect of negative (vs. neutral) valence, in which 6-year-olds responded marginally more accurately to negative than neutral words ($p = .06$).

3.2.2. *Language competence hypothesis*

As in the reaction time analyses, this set of analyses did not include 5-year olds. In addition the accuracy analyses excluded 7-year-olds. As such, our analysis of the language competence hypothesis in response accuracy only included data from 6-year-old children. The analysis included all previously mentioned control variables. Predictors of interest were language competence, valence, and their interaction. The analysis revealed that the interaction was not significant for negative (vs. neutral) valence ($p = .88$), nor for positive (vs. neutral) valence ($p = .82$).[3]

3.2.3. *Nimble-hands, nimble-minds*

We again included all previously mentioned control variables as well as Age, which served as a control variable in the present analysis. Predictors of interest were mean-centered BOI rating, BOT score, and their interaction. The analysis revealed that the interaction was not significant ($p = .72$), nor were the main effects of BOI ($p = .70$) or BOT ($p = .87$).[4]

---

[2] Note that we also ran this analysis only including five year-olds (the age group examined by Suggate & Stoeger, 2017) and also found no significant effects.

[3] Supplementary analyses using the first and second components found that they also did not interact with valence.

[4] Note that we also ran this analysis only including five year-olds (the age

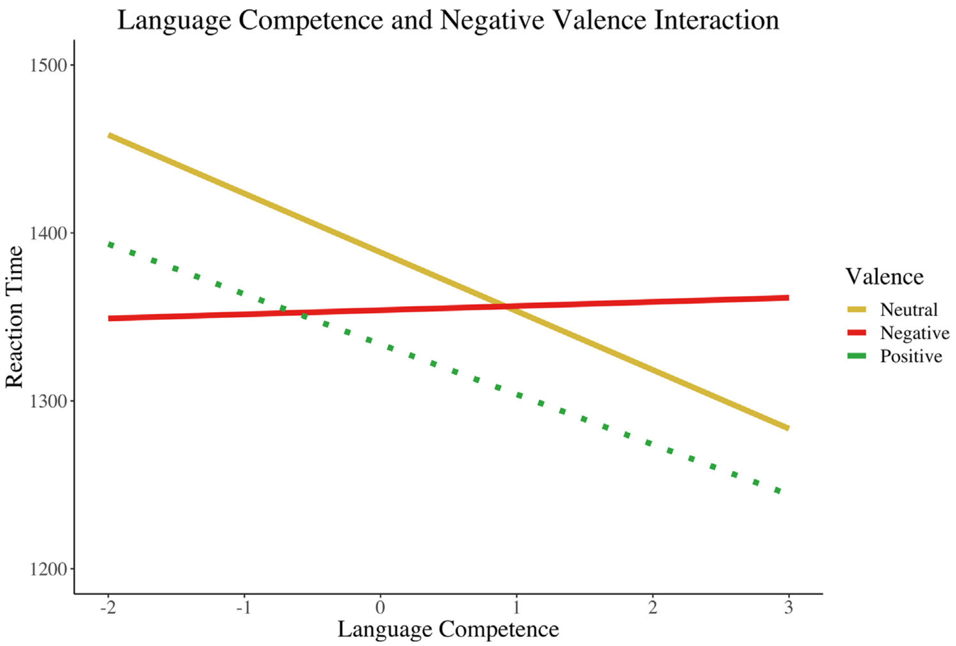


Fig. 4. Marginal plot for interaction between standardized language competence and valence in the prediction of reaction time. Positive (vs. neutral) valence also shown for reference, but in dashed form as not part of significant interaction.

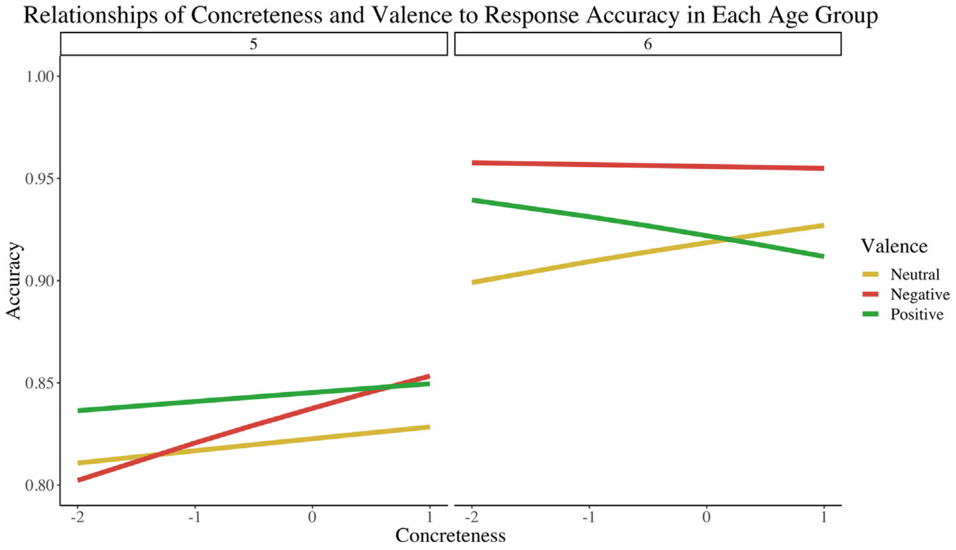


Fig. 5. Marginal plots for relationships of standardized concreteness and valence to response accuracy in each age group.

4. Discussion

The purpose of the present study was to test three proposals for vocabulary acquisition, derived from current theories of conceptual knowledge. The first proposal is that emotion provides a bootstrapping mechanism for vocabulary acquisition. Recent results (Ponari et al., 2017) suggested that around 8–9 years of age children show sensitivity to word valence in their ALDT responses. Interpreting the results for younger children in that study was complicated, however, by the low rate of ALDT accuracy among the 6–7 year olds tested. In the present study we also examined effects of valence on lexical processing, in younger groups of children. We used a large set of words that were, on average, acquired earlier than those presented in the previous study. With these more familiar words, children in the present study had higher accuracy in ALDT responses and we were able to examine children's reaction times. This is an important advance; reaction times are the primary source of evidence about underlying processes in the adult literature because they are less susceptible to floor and ceiling effects than are accuracy data.

We tested for valence effects in each of our 5-, 6-, and 7-year-old age groups and found that 6-year-old and 7-year-old children's ALDT reaction times were influenced by word valence. This sensitivity to emotion information was not present in the 5-year-old children we tested and was not significant in the accuracy analyses. Our reaction time findings provide evidence that at 6–7 years of age children ground word meanings via emotion systems (Kousta et al., 2011), suggesting that sensitivity to emotion information in lexical processing can be observed at a younger age than that inferred from previous research.

In the Ponari et al. (2017) study, the only significant valence effect was for abstract words in their intermediate (8–9 year old) age group,

---

(footnote continued)
group examined by Suggate & Stoeger, 2017) and also found no significant effects.

where accuracy was higher for positive words than for neutral words. There is little overlap between the stimuli used by Ponari et al. and those used in the present study (only 2 words in common), and our effects were observed in reaction times, yet we would argue that there is similarity in the basic pattern of sensitivity. In particular, we found faster reaction times for positive abstract words than for neutral abstract words in the present 6-year-old age group, and this is analogous to the accuracy advantage observed by Ponari et al. for positive abstract words vs neutral abstract words. While this pattern in similar across the two studies, we also found some differences. In the present study, 6-year-olds were faster (and tended to be more accurate, although not significantly so) for negative words than for neutral words. This did not interact with concreteness. This processing advantage for negative words is in keeping with some findings with adults: in visual lexical decision tasks adults have responded more quickly to negative words than neutral words (Kousta et al., 2009; Vinson et al., 2014; Yap & Seow, 2014). To our knowledge this has not previously been reported for child participants. In another departure from the Ponari et al. results, the 7-year-olds we tested showed faster responses to positively valenced words, and this too was not significantly modulated by concreteness.

One difference between the present study and that of Ponari et al. (2017) that may be important to explaining the different results is that the abstract items in Ponari et al. were numerically somewhat lower in mean concreteness rating ($M = 2.51$, $SD = 0.68$) than those we considered more abstract by our median split ($M = 2.77$, $SD = 0.69$). As such, the present manipulation of concreteness was likely weaker than that in the Ponari et al. study and this might have influenced the strength of the concreteness by valence interactions, rendering them nonsignificant in some cases. It is also true that the present participants were younger than those in the Ponari et al. study, so another explanation is that in this younger age group valence effects tend to be more generalized, whereas those observed in older children (e.g., the 8–9 year olds in Ponari et al.) tend to be more limited to abstract word stimuli. Future research will be required to adjudicate between these possibilities.

The valence effects that we observed in children's ALDT responses involved faster reaction times for positive and negative words than for neutral words. While this is consistent with many of the findings from the adult literature (e.g., Kousta et al., 2009; Siakaluk et al., 2016; Vinson et al., 2014; Yap & Seow, 2014), we noted in the Introduction that there is considerable variability in the particular ways in which valence influences adult lexical processing measures, and the reasons for this variability have not been established (Kuperman, 2015). It also seems possible that the particular effects of valence change across development, in the years after those tested in the present study. This will be an important issue for future research, but testing that issue will require different items than those used here. Since the current stimuli were selected for the age range tested (5–7 years), these items would be less suitable for older child and adult participants.

We also tested predictions of the language competence hypothesis, that language experience may be important to acquisition of abstract word meanings, particularly abstract neutral word meanings. Such words do not enjoy the benefit of valence information to ground meanings and must be learned via other mechanisms. We tested the prediction that language competence is important for acquisition of these word meanings once emotion information has already been recruited to capture differences between abstract and concrete word meanings (Vigliocco et al., 2018). Our results were consistent with this prediction: children with a greater degree of language competence responded faster to abstract items that were neutral in valence, but not those that were negative in valence. Further, our results suggest that the aspects of language competence that are related to processing of abstract neutral word meanings are those captured by the speech, syntax and coherence (i.e., producing easy to understand sentences) subscales of the CCC.

We also found an inhibitory relationship between language competence and processing of words of negative valence. Although highly speculative, we note that this could be consistent with more advanced language users beginning to transition to the alternative pattern that is sometimes observed for words of negative valence. That is, some adult studies have shown that responses are slower to negative words than to neutral words. As mentioned, the reasons for this alternative pattern are not well understood, and the pattern is usually attributed to vigilance to negative stimuli (Pratto & John, 1991). Our results suggest that one factor to consider in future research on valence effects for negative stimuli is language competence. We would suggest that related dimensions like executive function skills would also need to be considered, but the present findings may be useful as researchers work to understand the conditions under which effects of positive valence are distinct from effects of negative valence.

Finally, we tested predictions of the "nimble hands, nimble minds" proposal (Suggate & Stoeger, 2014, 2017) and found no evidence that children with better fine motor skills were better able to recruit sensorimotor information for word stimuli in the ALDT. The fact that we did not find support for the nimble hands, nimble minds hypothesis is problematic for a strong embodied account, as our findings suggests that embodied information is not always recruited in children's lexical processing. Although problematic for a strong embodied account (e.g., Glenberg & Gallese, 2012) our findings could be explained by a multimodal or multiple representations account, since those frameworks allow that there are multiple sources of information that support word knowledge (e.g., Borghi et al., 2017; Howell et al., 2005); embodied information is just one such source and its relevance to performance depends on the task and the type of concept.

4.1. Conclusions

In the present study we found that even 5-year-old children demonstrated reasonable rates of accuracy in the ALDT. As such, the task shows promise as a tool for exploring children's lexical-semantic development in future research, if the items used are familiar for children of the age tested. The results for the present version of the task provide new evidence that children recruit valence information in the process of word recognition, consistent with proposals like the Affective Embodiment Account (Borghi et al., 2017; Kousta et al., 2011) and also with broader proposals about grounded lexical development (Dove, 2011, 2018; Howell et al., 2005; Thill & Twomey, 2016) and multimodal semantic models (e.g., Andrews, Vigliocco, & Vinson, 2009; Barsalou et al., 2008; Borghi et al., 2017).

Declarations of interest

None.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada through a Discovery Grant to PMP (Grant number RGPIN/217309-2013). We thank Sophia van Hees for assistance with experiment programming.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2019.04.017.

References

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review, 116*, 463–498.
Bååth, R. (2010). ChildFreq: An online tool to explore word frequencies in child language.

LUCS Minor, 16, 1–6.

Baayen, H., Bates, D., Kliegle, R., & Vasishth, S. (2015). RePsychLing: Data sets from psychology and linguistics experiments. R package version 0.0.4.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283–316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445–459.

Barsalou, L. W., Santos, A., Simmons, K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.). *Symbols, embodiment and meaning* (pp. 245–284). Oxford, UK: Oxford University Press.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher, & R. A. Zwaan (Eds.). *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 129–163). Cambridge: Cambridge University Press.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv:1506.04967 [stat]. Retrieved from < http://arxiv.org/abs/1506.04967 > .

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015b). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.

Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*, 879–891.

Bishop, D. V. M. (2006). *CCC-2: Children's Communication Checklist-2, United States Edition, Manual.* San Antonio, TX: Pearson.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*, 341–345.

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin, 143*, 263–292.

Bruininks, R.H., & Bruininks, B. D. (2005). Bruininks-Oseretsky test of motor proficiency (2nd ed.).

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*, 904–911.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods Instruments & Computers, 36*, 384–387.

Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2010.00242.

Dove, G. (2018). Language as a disruptive technology: Abstract concepts, embodiment, and the flexible mind. *Philosophical Transactions of the Royal Society B, 373*, 20170135.

Dunn, L., & Dunn, D. (2007). *PPVT-4, Peabody picture vocabulary test manual* (4th ed.). Minneapolis, MN: Pearson Assessments.

Estes, Z., & Adelman, J. S. (2008). Automatic vigilance for negative words is categorical and general. *Emotion, 8*, 453–457.

Glenberg, A. M. (2015). Few believe the world is flat: How embodiment is changing the scientific understanding of cognition. *Canadian Journal of Experimental Psychology, 69*, 165–171.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex, 48*, 905–922.

Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M., & Steele, J. S. (2010). Fine motor skills and early comprehension of the world: Two new school readiness indicators. *Developmental Psychology, 46*, 1008–1017.

Howell, S. R., Jankowicz, D., & Becker, S. (2005). A model of grounded language acquisition: Sensorimotor features improve lexical and grammatical learning. *Journal of Memory & Language, 53*, 258–276.

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140*, 14–34.

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*, 473–481.

Kuperman, V. (2015). Virtual experiments in megastudies: A case study of language and emotion. *Quarterly Journal of Experimental Psychology, 68*, 1693–1710.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General, 143*, 1065–1081.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*, 978–990.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmer Test Package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26.

Long, J.A. (2018). jtools: Analysis and presentation of social scientific data. R Package Version 1.0.0.

Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics, 39*, 155–158.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers, 28*, 203–208.

Marshall, P. J. (2016). Embodiment and human development. *Child Development Perspectives, 10*, 245–250.

Moffat, M., Siakaluk, P. D., Sidhu, D. M., & Pexman, P. M. (2015). Situated conceptualization and semantic processing: Effects of emotional experience and context availability in semantic categorization and naming tasks. *Psychonomic Bulletin & Review, 22*, 408–419.

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Haremlen, A.-L., et al. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods, 45*, 169–177.

Pexman, P. M. (2012). Meaning based influences on visual word recognition. In J. S. Adelman (Vol. Ed.), *Visual word recognition: Vol. 2,* (pp. 24–43). New York, NY: Psychology Press.

Pexman, P. M. (2018). The role of embodiment in conceptual development. *Language, Cognition, & Neuroscience*. https://doi.org/10.1080/23273798.2017.1303522.

Pexman, P. M., & Wellsby, M. B. (2016). Linking hand to mouth: The relationship between manual dexterity and language skills in typically-developing children. *Italian Journal of Cognitive Sciences, 1*, 75–88.

Pexman, P. M., Muraki, E. J., Sidhu, D. M., Siakaluk, P. D., & Yap, M. J. (2019). Quantifying sensorimotor experience: Body-object interaction ratings for more than 9,000 English words. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-1171-z.

Ponari, M., Norbury, C. F., Rotaru, A., Lenci, A., & Vigliocco, G. (2018). Learning abstract words and concepts: Insights from developmental language disorder. *Philosophical Transactions of the Royal Society B, 373*, 20170140.

Ponari, M., Norbury, C. F., & Vigliocco, G. (2017). Acquisition of abstract concepts is influenced by emotional valence. *Developmental Science*. https://doi.org/10.1111/desc.12549.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*, 380–391.

R Core Team (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Schneider, W., Eschman, A., & Zuccolotto, A. (2001). *E-Prime user's guide.* Pittsburgh, PA: Psychology Software Tools.

Siakaluk, P. D., Newcombe, P. I., Duffels, B., Li, E., Sidhu, D. M., Yap, M. J., et al. (2016). Effects of emotional experience in lexical decision. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.01157.

Siakaluk, P. D., Pexman, P. M., Aguilera, L., Owen, W. J., & Sears, C. R. (2008). Evidence for the activation of sensorimotor information during visual word recognition: The body-object interaction effect. *Cognition, 106*, 433–443.

Singmann, H., Bolker, B., & Westfall, J. (2015). afex: Analysis of factorial experiments. R package version 0.23-0.

Suggate, S. P., & Stoeger, H. (2014). Do nimble hands make for nimble lexicons? Fine motor skills predict knowledge of embodied vocabulary items. *First Language, 34*, 244–261.

Suggate, S. P., & Stoeger, H. (2017). Fine motor skills enhance lexical processing of embodied vocabulary: A test of the nimble-hands, nimble-minds hypothesis. *The Quarterly Journal of Experimental Psychology, 70*, 2169–2187.

Thill, S., & Twomey, K. E. (2016). What's on the inside counts: A grounded account of concept acquisition and development. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.00402.

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition, 1*, 219–247.

Vigliocco, G., Ponari, M., & Norbury, C. F. (2017). The role of linguistic information in learning abstract words: Evidence from children with Specific Language Impairment (SLI). *Proceedings of the cognitive science society* (pp. 3441–3446). .

Vigliocco, G., Ponari, M., & Norbury, C. F. (2018). Learning and processing abstract words and concepts: Insights from typical and atypical development. *Topics in Cognitive Science, 10*, 533–549.

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion, 28*, 737–746.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavioral Research Methods, 45*, 1191–1207.

Wellsby, M., & Pexman, P. M. (2014). Developing embodied cognition: Insights from children's concepts and language processing. *Frontiers in Cognitive Science*. https://doi.org/10.3389/fpsyg.2014.00506.

Winkielman, P., Coulson, S., & Niedenthal, P. (2018). Dynamic grounding of emotion concepts. *Philosophical Transactions of the Royal Society B, 373*, 20170127.

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2012.00072.

Yap, M. J., & Seow, C. S. (2014). The influence of emotion on lexical processing: Insights from RT distributional analyses. *Psychonomic Bulletin & Review, 21*, 526–533.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*, 971–979.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The Educator's word frequency guide.* Brewster, NY: Touchstone Applied Science Associates Inc.