Acquisition of abstract concepts is influenced by emotional valence

Marta Ponari[a], Courtenay Frazier Norbury[b], Gabriella Vigliocco[c]

a. School of Psychology, University of Kent, Canterbury CT2 7NP, Kent, UK.

b. Royal Holloway, University of London, Egham, Surrey, UK.

c. Department of Experimental Psychology, University College London, 26 Bedford Way, London WC1H 0AP, UK.

Authors' note

Courtenay F. Norbury is now at Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PF, UK.

This research was supported by a grant from Nuffield Foundation.

Corresponding author: Marta Ponari, School of Psychology, University of Kent, Canterbury CT2 7NP, Kent, UK. Contact: m.ponari@kent.ac.uk

Research highlights

- Analysis of age-of-acquisition ratings suggests that the percentage of abstract words acquired increases following a sigmoid curve, with a steep increase around the age of 8-9.
- Analysis of the relation between emotional valence and concreteness suggests a role of valence in the acquisition of abstract words: both positive and negative abstract words are acquired earlier than negative.
- In an auditory lexical decision task with positive, negative and neutral abstract and concrete words, children aged 8-9 years show an advantage for positive abstract words compared to neutral.
- These findings provide constraints for theories of vocabulary acquisition and enable predictions for children with atypical language or emotional/social difficulties.

Abstract

There is considerable lack of evidence concerning the linguistic and cognitive skills underpinning abstract vocabulary acquisition. The present study considers the role of emotional valence in providing an embodied learning experience in which to anchor abstract meanings. First, analyses of adult ratings of age-of-acquisition, concreteness and valence demonstrate that abstract words acquired early tend to be emotionally valenced. Second, auditory Lexical Decision accuracies of children aged 6-7, 8-9, and 10-11 years (n = 20 per group) complement these analyses, demonstrating that emotional valence facilitates processing of abstract words, but not concrete. These findings provide the first evidence that young, school-aged children are sensitive to emotional valence and that this facilitates acquisition of abstract words.

The role of emotional valence in the acquisition of abstract concepts

The ability to mentally represent and use abstract concepts like *contempt*, *kindness*, *idea*, is uniquely human and underscores cultural, social and scientific development, yet our understanding of this remarkable ability is extremely limited. In adults' processing, concrete concepts have been considered easier to learn (e.g., Paivio, 1971), to remember (e.g., Begg & Paivio, 1969), and to process (James, 1975; Whaley, 1978; Rubin, 1980); an advantage that has been referred to as the "concreteness effect". In the past 30 years, various hypotheses have been advanced to explain differences in processing concrete and abstract concepts in healthy adults. For example, the Dual Code theory (Paivio 1971, 2007) suggests that concrete words are more easily accessed because they are represented in both an imagistic and verbal code, while abstract concepts are only represented in a verbal code. Another view is the 'context availability hypothesis' (Schwanenflugel, 1991), which argues that differences between concrete and abstract words come about within verbal semantic memory. Here, concrete concepts would have strong semantic associations with few contexts, whereas abstract words would have weak associations to a larger number of contexts. Thus, both theories assume that verbal information is central to the representation and processing of abstract words, either because language is the only format in which abstract words are represented (Paivio, 1971, 2007) or because any difference between concrete and abstract would be accounted for in terms of differences in semantic networks developed on the basis of linguistic information (Schwanenflugel, 1991).

Evidence from the imaging literature also indicates that abstract word representations are language-based; although processing abstract words appears to engage a more distributed network of brain regions than concrete words (Pexman, Hargreaves, Edwards, Henry & Goodyear, 2007), there is evidence that abstract word processing is associated with higher activation in the left inferior frontal gyrus and the superior temporo-lateral cortex (see Binder, Desai, Graves & Conant, 2009), two areas typically associated with the verbal semantic network.

Concreteness effects are also evident in child language; early studies of children's language production (Brown, 1957, as reported in Schwanenflugel, 1991) suggested that 75% of the words most frequently produced by school-age children (first to eight grade) were concrete, while only 28% of the most used words by adults were concrete. The few early studies investigating the learning of abstract and concrete words have yielded mixed results, often because of lack of control of lexical variables (e.g., Kiraly & Furlong, 1974; Richmond & McNinch, 1977). For example, McFalls, Schwanenflugel and Stahl (1996) reported concreteness effects on the accuracy (not on processing speed) of lexical decisions and naming in second grade children (controlling for grammatical class, frequency and length, but not for other variables such as familiarity and age of acquisition).

Concreteness effects in children are in line with a primary role of language skills in developing abstract vocabulary. For example, the 'syntactic bootstrapping' hypothesis (Gleitman, Cassidy, Nappa, Papafragou & Trueswell, 2005), argues that children can only learn abstract words after they achieve enough sophisticated linguistic knowledge to enable them to "match interpretations of ongoing events and states of affairs with the semantically relevant structures underlying co-occurring utterances" (p. 55). In other words, once the children have learnt a sufficient number of concrete words through 'word-to-world' mapping, they can start acquiring abstract words through a process of '(linguistic) structure-to-world' mapping. In this view, syntactic information has been argued to be crucial to this process as similarity in syntactic structure can indicate similarity in meaning (e.g., mental state verbs take a complement as in: "I think about…", I worry about…"). However, the process of bootstrapping from co-occurring utterances could be purely based on probabilistic co-occurrence of given words in given contexts, without necessarily a role for syntactic structure per se (Landauer & Dutnais, 1997; Andrews & Vigliocco, 2009). Either way, abstract words and concepts would be learnt via linguistic structure and context. We will refer to this view as the "Abstract-via-Language" (AvL) hypothesis.

It is however the case that abstract word processing is not always harder for adults. Kousta, Vigliocco, Vinson, Andrews and Del Campo (2011) reported that once all lexical factors contributing to speed of word recognition are controlled, abstract words are processed faster than concrete words. Critically, this study controlled variables such as imageability (i.e., how easily a mental image can be formed for a concept), age of acquisition and familiarity that have usually been confounded in previous studies. This reversal of the concreteness effect is explained on the basis of greater affective association of abstract, relative to concrete words. There is, in fact, a general tendency for abstract concepts to be more affectively loaded than concrete concepts, as indicated by a strong relation between concreteness ratings (judgements by speakers on the extent to which given words refer to concrete referents or not) and valence ratings (judgements on the extent to which given words have positive, negative or no emotional connotations) (Kousta et al., 2011; Vigliocco et al., 2014). Valence has been shown to modulate processing, i.e. emotionally valenced words are processed faster than neutral words (Kousta, Vinson & Vigliocco, 2009). Because abstract words tend to be emotionally valenced, once all other factors that favour concrete words (such as familiarity and imageability) are controlled, they are processed faster than concrete words (Kousta et al., 2011). Emotional information has also been shown to facilitate processing of abstract words (but not concrete) in naming (Moffat, Siakaluk, Sidhu & Pexman, 2015) and in semantic categorisation tasks (Moffat et al., 2015; Newcombe, Campbell, Siakaluk & Pexman, 2012), and to interfere with colour naming in the Stroop task (Siakaluk, Knol & Pexman, 2014). To complement these behavioural findings, Vigliocco et al. (2014) employed an fMRI study in which subjects performed lexical decisions for concrete and abstract words, again, critically controlling for lexical factors such as imageability and familiarity not controlled in other studies. Greater activation for abstract words was found in the rostral portion of anterior cingulate cortex (rACC), an area linked to amygdala and orbitofrontal cortex that is engaged in affective processing (Bush, Luu & Posner, 2000).

These findings led to the hypothesis that emotion provides grounding for abstract concepts (Kousta et al., 2011). The basic idea is that emotion, rather than language, may provide a bootstrapping mechanism for the development of abstract words and concepts: while words referring to concrete objects and actions would be learnt by associating sensory-motor experience with the word, abstract words would be learnt by associating emotional states with word. The association with emotional states could allow young children to set the ontological distinction between concepts grounded in the physical environment (concrete) and those grounded in our internal states (abstract), thereby bootstrapping the development of abstract knowledge. Traces of such link with the affective system would then be observed in processing during adulthood (Kousta et al., 2011; Vigliocco, Meteyard, Andrews & Kousta, 2009).

Kousta et al. (2011) present initial evidence in favour of this account, including findings from regression analyses showing how earlier acquired abstract words tend to have emotional connotations, which is a prerequisite for the hypothesis: if affective association of words supports the initial establishment of ontological distinction between referents in the world versus referents internal to the individual, then the first abstract words acquired should be emotionally valenced. We will refer to this view as the "Abstract-via-Emotion" (AvE) hypothesis. It is important to note that these two hypotheses (AvL and AvE) are not mutually exclusive and both types of processes are likely to contribute to the acquisition of abstract vocabulary. An open question is whether and to what extent they might differentially contribute at different developmental stages.

In general, we know very little about the development of abstract words, other than the observation that they are generally acquired later in development than concrete words. Two recent studies have investigated the development of emotion vocabulary in children aged 4-16 via parent or teacher report in English (Baron-Cohen, Golan, Wheelwright, Granader & Hill, 2010) and Chinese-speaking children aged 2-13 years (Li & Yu, 2015). Words included in these studies (336 in Baron-Cohen et al., and 363 in the Li & Yu study) "described mental state[s] with an emotional dimension" (Baron-Cohen et al., 2010, p. 2; examples were *furious* or *relief*), therefore they were all abstract words. Of particular interest from these studies is the observation that there is a sharp increase in the number of abstract words reportedly known by children between the ages of 7-8 and 9-10. However, these studies did not investigate growth in abstract word knowledge more broadly.

The present study aimed to provide initial insight into how abstract words are learnt and the role of valence in the processing of abstract and concrete words in early school years. First, we conducted a corpus analysis of ratings for over 13,000 English words provided by adult speakers. We assessed the relation between ratings of age-of-acquisition, concreteness and valence for this large sample of words, in order to estimate the number of abstract words thought to be learned during childhood and to further assess the extent to which the findings by Kousta et al. (2011) (based on a smaller sample of 1,500 words), generalise to this larger database. Finally, we conducted a cross-sectional auditory lexical decision study with children aged 6 to 11 years, in which we explore for the first time the processing of abstract and concrete words in an age range characterised by rapid and large increases in the number of abstract words acquired.

The relation between Concreteness, Valence and Age-of-Acquisition

1) Regression analyses of lexical databases

In order to obtain a first broad picture of the developmental trend in vocabulary acquisition of abstract words, we look at the relation between concreteness (Brysbaert, Warriner & Kuperman, 2014) and age-of-acquisition (Kuperman, Stadthagen-Gonzalez & Brysbaert, 2012) ratings for more than 13,000 words, obtained from adult native English speakers. In addition, in order to establish whether earlier acquired abstract words tend to be more emotionally valenced than concrete words, we considered the relation between age-of-acquisition and valence (Warriner, Kuperman & Brysbaert, 2013) ratings, separately for concrete and abstract words.

Method

Words and databases used

The study employed a set of 13,266 words for which ratings for age-of-acquisition (age at which given words are learnt), concreteness (the extent to which given words refer to concrete referents or not), and valence (the extent to which given words have positive, negative or neutral emotional connotations), were available from existing databases. AoA ratings were taken from Kuperman et al. (2012), who asked adult participants to enter the age (in years) at which they thought they had learnt a word; age-of-acquisition ratings ranged from 2.7 to 19.21. Although these subjective ratings are estimates by adult speakers of the age at which words are learnt, they are highly correlated with recent measures of word frequency for primary school children taken from the CBBC channel's subtitles (a BBC channel targeted at children between 6 and 12 years of age; van Heuven, Mandera, Keuleers & Brysbaert, 2014): CBBC frequency is a significant predictor of AoA ratings, $F(1, 13245) = 9566.4$, $p < .001$, with lower CBBC frequencies corresponding to higher AoA ratings.

Concreteness ratings were collected by Brysbaert et al. (2014), using on a 5-point rating scale ranging from 1 = very abstract ("language based") to 5 = very concrete ("experience based"); we considered as concrete all the words in the database with concreteness ratings ≥ 3, and as abstract, those words with concreteness ratings < 3. In Brysbaert et al. (2014), instructions emphasised the distinction between words that refer to actions and objects one can have immediate experience of with the senses (concrete), and those that do not refer to something one can experience with their senses or actions but that depend on language. Valence ratings were taken from Warriner et al. (2013; following the ANEW database, Bradley & Lang, 1999; and also Kousta et al., 2011); they were collected on a 1 to 9 scale, with 1 = very negative ("completely unhappy"), 5 = neutral and 9 = very positive ("completely happy").

<u>Proportion of abstract vocabulary at different ages</u>

As expected, inspection of AoA ratings provided by adults confirms that abstract words are rated to be acquired later than concrete words. As illustrated in Figure 1, for children under the age of four, AoA ratings suggest that less than 10% of vocabulary is abstract; abstract vocabulary then is thought to increase dramatically to more than 40% by the age of 12. In order to estimate how abstract vocabulary changes during childhood, words were ordered by age-of-acquisition (from low to high) and sampled to calculate the percentage of abstract words thought to be known at 0.1-year intervals from 2.5 to 19, thus obtaining 151 data points. These data were used to fit a polynomial regression model, using R (version 3.0.2; R Core Team, 2013). The best fitting function was a third-degree polynomial (1), with an adjusted R-squared of 0.987, indicative of an extremely high level of fit.

$$f(x) = -0.029x^3 + 0.0751x^2 - 2.618x + 10.762 \quad (1)$$

The second derivative of the values predicted by the function was calculated to determine the inflection point (i.e., the point at which the curve changes from being convex to concave), which was between 8.5 and 8.6 years of age. This indicates that, according to age-of-acquisition ratings, the increase in the percentage of abstract words known is steeper around the age of 8.5 years.

~ Please insert Figure 1 about here ~

<u>Relations between AoA, concreteness and valence</u>

In a first linear regression analysis, we looked at the relation between valence and concreteness ratings. Both linear and non-linear components of valence were significant predictor of concreteness ratings, $F(2, 13111) = 209.02$, $p < .0001$; nonlinear $F(1, 13111) = 287.19$, $p < .0001$. This confirms and extends, using a much larger set of words, Vigliocco et al. (2014) conclusion that the more valenced a word is, the more likely it is to be abstract; conversely, more neutral words tend to be concrete (see Figure 2).

~ Please insert Figure 2 about here ~

We carried out a second set of regression analyses separately on abstract (5,461) and concrete (7,654) words using polynomial models in order to formally assess the relation between AoA ratings and valence for abstract and concrete words.

We found that for both abstract and concrete words, valence and AoA ratings are related by a quadratic function (combined linear and quadratic components for abstract words: $F(2, 5458) = 289.75$, $p < .0001$; quadratic alone: $F(1, 5458) = 420.97$, $p < .0001$; combined linear and quadratic components for concrete words: $F(2, 7651) = 435.00$, $p < .0001$; quadratic alone: $F(1, 7651) = 166.72$, $p < .0001$). The shape of the function differs, however; to better interpret the different functions, we divided abstract and concrete words into valence categories (negative: valence ratings < 4.0; positive: valence ratings > 6.0; and neutral: valence ratings of 4.5-5.5) and computed t-tests. For abstract words (see Figure 3, left panel), AoA ratings suggest that both positive and negative words are acquired earlier than neutral (positive vs neutral: $t(2700) = 18.005$, $p < .001$; negative vs neutral: $t(2816) = -6.156$, $p < .001$). For concrete words (see Figure 3, right panel), instead, neutral words seem to be acquired earlier than negative ($t(3961) = 5.669$, $p < .001$), while positive words are acquired earlier than both negative ($t(2970) = 22.820$, $p < .001$) and neutral ($t(4341) = 21.010$, $p < .001$).

~ Please insert Figure 3 about here ~

Discussion

As a first step in elucidating the acquisition of abstract words in childhood, we considered the proportion of abstract vocabulary that is rated as being acquired at different ages, using data from adult ratings of AoA and concreteness. Here we document for the first time the development of abstract vocabulary in childhood, from being less than 10% of the total vocabulary of children aged 4, to more than 40% by the age of 12. Especially interesting is the observation that the rate at which abstract vocabulary is acquired appears to change at about 8.5 years of age. Consistent with the strong correlation between emotion and concreteness (Kousta et al., 2011; Vigliocco et al., 2014), the steep increase in abstract word knowledge up to an age of 8.5 resembles the sharp increase in knowledge of emotion words reported by Baron-Cohen et al. (2010, for British children) and by Li and Yu (2015, for Chinese children) up to 9 years of age.

Second, we assessed the relation between AoA ratings, concreteness and valence. We found that for abstract words, valenced words (both positive and negative) appear to be learnt earlier than neutral words. For concrete words, we observed that positive words are also learnt early, whereas negative words appear to be acquired slightly later than neutral words. These findings are consistent with a general statistical relation between valence and concreteness (with more abstract words being valenced). This link to emotion facilitates language processing of abstract words in adults (Kousta et al., 2009; 2011). The finding that the first abstract words acquired are valenced (both positive and negative) provides support for the hypothesis that emotion may provide a bootstrapping mechanism for the acquisition of abstract words. As described in the introduction, the engagement of the affective system may be critical in enabling children to learn that some words refer to internal states, rather than to objects and actions (Kousta et al., 2011; Vigliocco et al., 2014). Importantly, this is not to say that only abstract words are linked to valence, concrete words can be valenced too; and, in fact, we also observed that for concrete words, positive words are learnt first. For concrete words, however, emotion may not be such a reliable of a cue to word meaning, given that for these words, many referents exist in the physical world.

It is interesting to see that for both concrete and abstract words, there seem to be an advantage for positive words. The reasons for this positive word bias is unclear; however, we may speculate that it represent a bias, at least in Western countries, toward using positive language with children. To support this, we have looked at the Child Language Data Exchange System (CHILDES) subcorpus of child-directed speech (MacWhinney, 2000). Of the 766 words (out of 1512) for which we have both valence (from Warriner et al. 2013) and concrete ratings (from Brysbaert et al., 2014), 363 are positive and only 63 are negative. In particular, of the 162 abstract words, more than 70% are valenced (89 positive, 25 negative); of the 604 concrete words, 52% are valenced (274 positive, 38 negative). Of the 50 most frequent words, more than half are positive and none are negative. This tendency to use positive (or to avoid negative) words in child-directed speech is plausibly the source of the positive bias we see stronger in AoA ratings, but also in adults' lexical decisions (Kuperman, Estes, Brysbaert & Warriner, 2014). Similar support for a positive bias can be found in the Oxford Communicative Development Inventory (OCDI; Hamilton, Plunkett & Schafer, 2000). The OCDI is a list of 433 words that are considered common in children's vocabulary, and it is widely used as a tool for measuring vocabulary size and growth. Of the 345 words for which we have both concreteness and valence ratings, only less than 10% are abstract (34 words); 68% of the 34 abstract words are emotionally valenced, with a strong positive bias (6 negative, 17 positive). Of the 311 concrete words, 56% are valenced, again with a strong positive bias (15 negative, 159 positive). Thus, data from the CHILDES corpus of child-directed speech and of the OCDI words, in line with our analyses of AoA ratings, show that abstract words are learnt later than concrete (as indicated by the small number of abstract words present in these databases), and also that more abstract than concrete words are emotionally valenced, particularly positive, although there is a positive bias for concrete words.

While these analyses provide first evidence concerning the type of abstract (and concrete) words known by children at different ages, they are nonetheless based on AoA ratings provided by adults. As we have shown that these ratings highly correlate with age-relevant frequencies (extracted from subtitles from children's BBC channels CBBC and CBeebies), and a recent paper has also independently validated the use of these ratings in age of acquisition research (see Brysbaert, *in press*), these ratings provide a good estimate of children's knowledge, or at least of children's exposure to words (the higher the AoA rating, the lower the frequency of occurrence in TV shows aimed at children; and vice-versa). However, to know when children are exposed to specific abstract and concrete words does not necessarily imply that the same regularities will be found in children knowledge of these words, thus converging evidence is important. In the following experiment, the first to directly assess knowledge of abstract and concrete words across the age range (6-12) corresponding to a steep increase in abstract knowledge, we directly assess processing using an auditory lexical decision task.

<u>Processing abstract and concrete words children:</u>

<u>A lexical decision experiment</u>

This study assesses processing of abstract and concrete words in children aged 6 to 12. We chose to use a lexical decision task in keeping with the literature on adult processing (e.g., Kousta et al., 2011), most of which has employed this paradigm to tap the earliest stages of processing; however, to reduce the effect of individual reading abilities, here we presented the words auditorily. We selected positive, negative and neutral abstract and concrete words controlling for variables that we know affect recognition accuracy in adults, including length (number of letters), frequency and AoA (from adult ratings). Controlling for these variables that typically favour concrete words should reduce the concreteness effect, as it was also previously reported with adult participants (e.g., Kousta et al., 2011).

Of special interest is therefore to see: (i) whether and to what extent abstract valenced words are recognised more accurately than abstract neutral words, as we would expect on the basis of the AvE hypothesis; (ii) whether any advantage for abstract valenced words is stronger for younger than older children, as suggested by AvE but not by AvL, as the emotional grounding of abstract words would be especially important at earlier stages of acquisition.

Method

Participants

Sixty typically developing children aged 6-12 years (26 males; mean age = 8.09, SD = 1.85; range = 6.08-11.79) were recruited from mainstream classrooms in Southeast England. Children were divided in three age groups: 6-7 years; 8-9 years; 10-11 years. All participants were native English speakers, had normal or corrected to normal vision and no history of developmental disorders or reported special educational need. The study was approved by the University College London Research Ethics

Committee. Informed, written consent was obtained from the parents, and verbal assent was obtained from the children prior to testing.

Background assessment

Children's non-verbal cognitive abilities were assessed using the Matrix Reasoning test of the Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler, 1999). The British Picture Vocabulary Scale (BPVS: Dunn, Dunn, Whetton, & Burley, 1997) was used to obtain an objective measure of receptive vocabulary.

Five children were not included in analyses due to poor performance on one or both of the standardised tests (standard score lower than 80 on the BPVS; t-score of 30 or less matrix reasoning test), leaving 55 participants for further analysis (19 children in group age A, 18 in groups B and C).

Materials

Twenty-four abstract and 24 concrete words were selected from a pool of 3,505 words for which normative data on a range of lexical variables could be obtained. These variables included ratings of AoA (Kuperman et al., 2012), concreteness, familiarity (Coltheart, 1981), valence (Warriner et al., 2013), and (log)frequency (Balota et al., 2007). AoA ratings were used to divide words into AoA bands (1: words acquired at 4-5 years; 2: words acquired at 6-7 years; 3: words acquired at 8-9 years; 4: words acquired at 10-11 years). Within each AoA band, triplets of negative (valence ratings < 4.0), positive (valence ratings > 6.0) and neutral (valence ratings of 4.5-5.5) words matched on length (number of letters), concreteness and (log)frequency were created. Each abstract triplet was then paired to a concrete triplet matching for length and frequency. Adults' lexical decision accuracy and reaction times (taken from the English Lexicon Project, Balota et al., 2007) did not vary between abstract and concrete words. Concrete and abstract sets differed in imageability (with concrete words being, on average, more imageable than abstract), but within concrete and abstract words there was no difference in imageability between valence categories. Concrete and abstract sets did not differ on familiarity, and on a measure of frequency taken from subtitles from a UK TV channel targeted at children aged 6-12 (CBBC; Van Heuven et al., 2014). Lexical and sublexical characteristics of the words are listed in Table 1, and a list of all 48 words used is presented in Appendix A.

Forty-eight pronounceable nonwords were created by changing one phoneme from 48 words matched to the experimental words on length, AoA, valence and concreteness. Words and nonwords were recorded by a native English speaker using Audacity v. 1.2.2.

~ Please insert Table 1 about here ~

Procedure

All children were assessed in a quiet room in their school and received stickers for participation. Words and nonwords were presented acoustically using E-Prime version 2.0 software (Psychology Software Tools, Pittsburgh, PA) running on a Dell Latitude E6320 laptop with a touchscreen display. Participants sat a comfortable distance from the computer screen and were presented with a computer game in which they were asked to help a cartoon alien learn English. After receiving verbal instructions, children were asked to wear headphones and were then presented with six practice trials (three nonwords and three words not used in the experiment). During the practice trials, visual feedback was given after each trial through images of either a smiling (correct trial) or frowning (incorrect) cartoon alien.

~ Please insert Figure 4 about here ~

In each trial, a cartoon alien was presented in the middle of the screen for 1000 ms, followed by the auditory presentation of either a real English word or a nonword. Immediately after the offset of the word (average stimulus duration = 830 ms), two touch screen buttons appeared at the bottom left (a red thumbs-down icon) or the bottom right (a green thumbs-up icon) of the screen (see Figure 4). Children were instructed to press the green button when they heard a word they knew, or the red button if they heard a "funny made-up" word. After the six practice items, participants completed all 96 items (24 abstract and 24 concrete words, plus 48 nonwords) presented in a randomised order. To minimise fatigue, children were given the choice to take a break after each block of 24 trials. Accuracy and reaction times were recorded.

Design and Data Analysis

We focus on accuracy (proportion correct) because, given the high constraints in selecting stimuli, we could not control for first phonemes or for uniqueness point, both of which are known to affect RTs in auditory lexical decision tasks (see Goldinger, 1996). In order to exclude children who show a bias toward either always answering "word" or "nonword", we computed the rate of *hits* (correctly identifying a word) and *false alarms* (incorrectly claiming that a nonword was a word). Although all children, regardless of their age, saw all the 48 words (thus for the younger children, including words from higher AoA bands that they would not be expected to know), *hits* were only computed for words in the age-relevant AoA bands. We computed the response bias (or *criterion, c*), calculated by multiplying the sum of the normalised *hit* rate and the normalised *false alarm* rate by -0.5 (Fox, 2004; MacMillan & Creelman, 1991; Shapiro, 1994). A criterion with a negative value would indicate that responses are biased toward answering "word" (both words and nonwords are more likely to be indicated as words); a criterion of positive value would, conversely, indicate a response bias toward answering "nonword" (both words and nonwords are more likely to be indicated as nonwords). The average criterion bias was 0.02 (SD = 0.33) for 6-7 years old, -0.06 (SD = 0.45) for 8 - 9 years old, 0.05 (SD = 0.31) for 10-11 years old. Children who showed a criterion bias higher than 1.5 standard deviations above their group mean (indicating a strong bias toward "nonword" responses) or lower than 1.5 standard deviations below their group mean (indicating a strong bias toward "word" responses) were excluded from further analyses. Using these criteria, we excluded 1 child in the 6-7 age group (c = 0.48); two children in the 8-9 age group (c_1 = -1.12, c_2 = 0.72); and one child in 10-11 age group (c = 0.75). This left us with 18 6-7 years old children, 16 8-9 years old children and 17 10-11 years old children for further analyses (see Table 2).

~ Please insert Table 2 about here ~

Accuracy rates were analysed using mixed logit models (LME; package 'lmerTest' (Kuznetsova, Brockhoff & Christensen, 2016), running in R version 3.2.1 (R Core Team, 2015). We started by fitting baseline models that included factors that were not matched triplet-wise in the stimuli: familiarity, imageability (from Coltheart, 1981), and CBBC frequency (van Heuven et al., 2014), along with our categorical variables of interest: valence (positive, negative, neutral), concreteness (abstract, concrete) and age group (6-7, 8-9, 10-11), as well as the three-way interaction between these. All continuous predictors were centered on the mean. Whenever possible, the maximal by-child and by-item random structure was included (Barr, Levy, Scheepers & Tily, 2013). Log-likelihood ratio tests were used to compare fitted models.

Results

The baseline model included all predictors mentioned in the Data Analyses section above, and the three-way interaction between concreteness, valence and age group. First, we compared the baseline model against a model that included all combinations of two-way interactions and the main effects. Including the three-way interaction significantly improved the fit of the model (log-likelihood ratio for interaction model = -759.43; log-likelihood ratio for two-way interactions model = -765.46; $\chi^2(4) = 12.046$, p = .017). In this model, familiarity (coefficient estimate = 0.49, SE = 0.20, p = .022) and CBBC frequency (coeff. estimate = 0.75, SE = 0.15, p < .001) were significant predictors of children's performance, so they were kept in subsequent models. There was no significant main effect of Concreteness (coefficient estimate = -0.10, SE = 0.07, p = .173). To better explore the three-way interaction (see Figure 5), we fitted further models looking at the interaction between valence and age group, separately for abstract and concrete words.

~ Please insert Figure 5 about here ~

Abstract words

Including the valence × age group interaction marginally improved the fit of the model (log-likelihood ratio for interaction model = -430.36; log-likelihood ratio for main effects model = -434.66; $\chi^2(4) = 8.607$, p = .072). We fitted further models to explore this interaction; to allow for clearer interpretation of the parameter estimates, we used two-level factors (along with the same control variables as in the final model above). Therefore, a) we contrasted the performance of age groups pairwise (6-7 vs. 8-9 years old; 8-9 vs. 10-11 years old); and b) we contrasted the three valence categories pairwise (positive vs. neutral; negative vs. neutral; positive vs. negative). The fixed effects of interest were contrast-coded to ensure that interactions and main effects were orthogonal.

6-7 vs 8-9 years old children

Positive vs. Neutral: The main effect of valence was not significant (estimate of the Positive-Neutral difference = 0.075, SE = 0.05, t = 1.676, p = .113), but there was a significant main effect of age group (estimate of the 8-9 vs 6-7 years old difference = 0.190, SE = 0.04, t = 4.351, p < .001). The valence × age group interaction was marginally significant (coefficient estimate[1] = 0.118, SE = 0.06, t = 1.920, p = .055). Further models revealed that, although 8-9 years old children were better than 6-7 with both positive (t = 5.052, p < .001) and neutral (t = 2.139, p = .040) words, the difference between positive and neutral words was not significant in 6-7 years old children, t = 0.232, p = .819; while it was in 8-9 years old, t = 2.389, p = .030, with positive words being recognised better than neutral.

Negative vs. Neutral: There was no significant main effect of valence (estimate of the Negative-Neutral difference = -0.013, SE = 0.08, t = -0.177, p = .861), and no significant valence × age group interaction (coefficient estimate = 0.050, SE = 0.06, t = 0.808, p = .420). The main effect of age group was significant (estimate of the 8-9 vs 6-7 years old difference = 0.156, SE = 0.05, t = 3.388, p = .002).

Positive vs Negative: there was no significant main effect of valence (estimate of the Positive-Negative difference = 0.111, SE = 0.06, t = 1.727, p = .103) and no valence × age group interaction (coefficient estimate = 0.066, SE = 0.06, t = 1.165, p = .244).

[1] In the case of interactions, ‘coefficient estimate’ refers to the difference of difference scores. In this case it refers to the ‘8-9YearsOld(Positive-Neutral) - 6-7YearOld(Positive-Neutral)’ difference, thus testing whether the magnitude of the valence effect differs for 6-7 years old and 8-9 years old children.

The main effect of age group was significant (estimate of the 8-9 vs 6-7 years old difference = 0.215, SE = 0.04, t = 5.079, p < .001).

8-9 vs 10-11 years old children

Positive vs. Neutral: The main effect of valence was not significant (estimate of the Positive-Neutral difference = 0.091, SE = 0.05, t = 1.654, p = .118), nor was the main effect of age group (estimate of the 10-11 vs. 8-9 years old difference = 0.054, SE = 0.03, t = 1.620, p = .115); however, the valence × age group interaction was significant (coefficient estimate = -0.143, SE = 0.05, t = -2.822, p = .005): the difference between positive and neutral words, significant at 8-9 years old (see above), was not significant for 10-11 years old children, t = 0.621, p = .543. Furthermore, 10-11 years old children were better than 8-9 years old with neutral words (t = 2.284, p = .018), but not positive (t = -0.499, p = .621).

Negative vs. Neutral: There was no significant main effect of valence (estimate of the Negative-Neutral difference = -0.038, SE = 0.10, t = -0.381, p = .708). The main effect of age group was marginally significant (estimate of the 10-11 vs 8-9 years old difference = 0.064, SE = 0.03, t = 1.894, p = .067). The valence × age group interaction was also significant (coefficient estimate = -0.122, SE = 0.05, t = -2.255, p = .025). The difference between negative and neutral words was not significant in both groups, but once again the 10-11 years old children were better than the 8-9 years old with neutral words, but not with negative (t = 0.092, p = .927).

Positive vs Negative: there was no main effect of valence (estimate of the Positive-Negative difference = 0.116, SE = 0.09, t = 1.295, p = .214), no significant main effect of age group (estimate of the 10-11 vs 8-9 years old difference = -0.007, SE

= 0.03, t = -0.245, p = .808), and no valence × age group interaction (coefficient estimate = -0.021, SE = 0.04, t = -0.461, p = .645).

Concrete words

Including the valence × age group interaction did not improve the fit of the model (log-likelihood ratio for interaction model = -330.69; log-likelihood ratio for main effects model = -333.25; $\chi^2(4) = 5.111$, p = .276). To test the main effects, we proceed by removing them one at a time: removing the effect of valence marginally reduced the fit (log-likelihood ratio for the model including valence = -333.25; log-likelihood ratio for the model without = -335.98; $\chi^2(2) = 5.474$, p = .065). To interpret the main effect of valence, we contrasted valence categories pairwise: neutral words were recognised marginally better than both negative (coefficient estimate = -0.133, SE = 0.07, t = -1.968, p = .067) and positive (coefficient estimate = -0.094, SE = 0.05, t = -1.854, p = .084); no difference was found between positive and negative (t = 1.256, p = .228). Removing the main effect of age also significantly reduced the fit (log-likelihood ratio for the model including age group = -333.25; log-likelihood ratio for the model without = -354.26; $\chi^2(2) = 42.018$, p < .001). Recognition of concrete words was better at 8-9 years old compared to 6-7 (estimate of the 8-9 vs 6-7 years old difference = 0.169, SE = 0.04, t = 4.375, p < .001) and at 10-11 years old compared to 8-9 (estimate of the 10-11 vs 8-9 years old difference = 0.082, SE = 0.02, t = 4.146, p < .001).

Discussion

In this study we tracked for the first time children's knowledge of abstract words from age 6 to age 12. Overall, we found that, for both abstract and concrete words, accuracy increased with age; we also found that valence interacted with age for abstract but not for concrete words. No overall effect of concreteness was observed, perhaps because of the tight control of those lexical variables (such as length and familiarity) that favour concrete words, as also found in adults' studies (Kousta et al., 2011). The significant interaction between valence and age indicates that valence affected accuracy rates for abstract words differently at different ages. In particular, we found that children aged 8-9 were the only group showing an effect of valence; particularly, they were better with positive abstract words compared to neutral words. This suggests that children aged 8-9 years old particularly benefit from word valence (especially positive) when learning abstract words; later (at 10-11), as more words are learned, knowledge of neutral abstract words increases, and children do not show an advantage for valenced abstract words anymore. The lack of valence effect for the younger children (aged 6-7) is more difficult to interpret, as their overall accuracy is very low; in fact, only half of the items in the experiment have AoA $\leqq$ 7 (see methods), thus raising the possibility of a floor effect. This is not the case for the two other groups for which accuracy is high and for whom there are more known items.

The fact that the valence effect we observed is actually a positive effect is in line with the analyses of AoA ratings reported above, and also the follow-up analyses of the early acquired words (from OCDI and CHILDES); this might reflect a tendency to avoid negative words in the language to which children are exposed, especially earlier on.

For concrete words, we found an advantage for neutral words across the three age groups. This advantage may reflect different strategies used to learn concrete and abstract words, as suggested by Kousta et al (2011). The idea is that in learning concrete words (which, statistically, tend to be neutral) children focus on their sensory-motor properties, not their affective properties. Children would, instead, focus on affective properties for abstract words. Given the use of such strategies, other things being equal, concrete emotional words (as well as abstract neutral words) could be somewhat disadvantaged as they are less typical. However, clearly this is not the whole story as our analysis of AoA indicates that concrete positive words are also learnt early.

General Discussion

In this study we have investigated for the first time the developmental trajectory of learning abstract words by combining data from lexical databases obtained from adult speakers and empirical data from an auditory lexical decision task in a cross-sectional study of children aged 6-12 years. The results from both studies converge and complement each other. In both studies, we have established a significant role for valence, especially positive, in the acquisition of abstract vocabulary. In the analyses of the relationship between AoA ratings and valence, we found that both positive and negative abstract words are acquired earlier than neutral ones, however it also seems that positive words are acquired earlier than negative. In the lexical decision study, we found that 8-9 years old show an advantage for positive abstract words compared to neutral; this advantage disappears as knowledge of neutral words increases at 10-11 years of age. Crucially, it is not present for concrete words.

This is also in line with the literature on adult processing (lexical decision task), in which the picture emerging suggests a potential difference between positive and negative words. For example, Kousta et al. (2009) reported an advantage in accuracy for both positive and negative words compared to neutral, but the effect was stronger for positive words. Kuperman et al. (2014) also found that the advantage for emotional words was larger for positive than for negative words (and modulated by frequency).

These differences in adult processing could be rooted in differences in the age of acquisition of positive and negative words. Our analyses of the database indicate that the advantage for positive words is not specific for abstract words but also encompasses concrete words. In fact, according to our database analyses, positively valenced concrete words tend to be acquired earlier than other concrete words. A possible explanation for this positive advantage is the prevalence of positive emotional language to which children are exposed in child directed speech.

Moreover, in our lexical database analyses, we established that the percentage of abstract vocabulary steeply increases up to around 8-9 years of age, after which it flattens down. Ages 8-9 years is also the age at which, in our lexical decision study, children seem to benefit the most from positive valence. Interestingly, a trend, similar to the one we observe in our analysis of lexical databases, is also reported in the study of Baron-Cohen et al (2010) for emotional words. Altogether, these findings present a nuanced but also very interesting picture concerning the role of valence in learning abstract words. Implications for the theoretical perspectives presented in the introduction are discussed below.

In the introduction, we described two main theoretical views concerning how abstract words could be learnt. According to AvL, learning abstract words and concepts would be entirely based on language. One potential mechanism underlying language-based vocabulary acquisition, that would benefit especially abstract words, would be the extraction of correlational patterns in discourse and texts. On the basis of the linguistic contexts in which a word is used, children could make inferences about their meaning, as argued by distributional theories of semantics (Landauer & Dutnais, 1997; Griffiths, Steyvers & Tenenbaum, 2007; Andrews & Vigliocco, 2009). A related hypothesis is the "syntactic bootstrapping" hypothesis according to which it is specifically syntactic information that would support the learning of abstract words, especially those that denote events and states (Gleitman et al., 2005). Of course, in order for such mechanisms to be effective, children must have acquired ample linguistic competence and vocabulary. Here, it is important to note that the ages we are investigating also correspond to the time in which children become skilled readers and therefore they can greatly expand their vocabularies taking advantage of co-occurences in texts. Language is further used in an explicit manner when carers or teachers tell children the definition of abstract words. While we are not debating whether these mechanisms are used, we question whether this could be the only manner in which children learn abstract words.

The AvE hypothesis concerns the early stages of learning abstract words, it provides a mechanism for bootstrapping the learning of abstract concepts and words that takes advantage of the strong association between abstractness and emotional valence. AvE argues that emotional valence could support the establishment of the distinction between concrete and abstract domains of knowledge; while concrete words would refer to observable entities and actions that we can experience with our senses and act upon, abstract words would refer to internal states of self and others that trigger embodied emotional reactions and experiences. These emotional reactions could come about from interactions with caregivers in which children would associate words being heard with emotions being expressed by the caregivers or by the child themselves. In such a view, the communicative social interaction would play a central role, along the lines proposed by recent social-cognitive theories of lexical development (e.g., Tomasello, 2010).

What do our results tell us about these theories? As predicted by AvE, analyses of AoA ratings show that emotionally valenced abstract words are acquired earlier than neutral, thus allowing emotion to bootstrap abstract knowledge. Moreover, results of the auditory lexical decision study show that around the age of 8-9 years old, children are more accurate in processing abstract positive, relative to neutral abstract words. The picture, however, is nuanced as it is especially positive affect that underscores the relationship between AoA and valence (across both concrete and abstract words). Still for abstract, and not for concrete, both positive and negative words have earlier AoA ratings. The positive bias is then reflected in processing accuracy for abstract, but not for concrete, words in our experiment (for children aged 8-9). One may ask why we did not find higher accuracy for abstract positive words for the younger (6-7) children as predicted by the AvE. One possible reason for this is that accuracy was particularly low in this group, given that only half of our 48 words had age-appropriate AoA ratings, and therefore any small effect could be masked by an overall floor effect.

The number of abstract neutral words increase greatly after the age of 10, according to our AoA analysis and our lexical decision experiment shows that for children aged 10-11, there is no more any advantage for positive abstract words. It is likely that by this age, strategies that are not grounded on emotion are being used more. It is also interesting to note that this is also the age at which children know similar percentages of abstract and concrete words, as shown in our analysis of lexical databases, which might suggest that emotional valence is not as relevant after the age of 9. As argued by the AvE, emotion might be especially relevant early on, but once vocabulary increases, valence alone does not allow for fine grade discrimination between abstract words with similar meaning. At the same time, with a wider vocabulary and linguistic competence, the child can use more effectively the correlational patterns in discourse in order to extrapolate abstract meaning. As reading proficiency develops, children are also more likely to acquire new vocabulary from written texts, rather than through interpersonal exchanges, and recover abstract meaning from surrounding linguistic context.

Conclusions

We have reported converging evidence from both analyses of age of acquisition ratings and children's performance on an auditory lexical decision task, showing that emotion (especially positive) plays a role in learning early abstract words particularly at around the age of 8-9, while language may play a greater role for abstract words acquired after this age. In addition to informing theories of how children acquire abstract knowledge, these results provide the first normative data on the acquisition of abstract vocabulary and enable predictions for children with atypical language or emotional and social development. Previous research shows that children with specific language impairment (SLI) have difficulties using syntactic bootstrapping to learn new words (Shulman & Guberman, 2007), which according to AvL would predict a marked deficit in abstract vocabulary, relative to other types of abstract knowledge. Our results suggest that children with SLI would not be impaired in the acquisition of abstract concepts that are generally acquired before the age of 8-9, as vocabulary at this age could be supported by emotional valence. On the other hand, children with social or emotional difficulties might be initially impaired, even if their linguistic abilities are on track. Further studies using clinical populations should investigate these predictions.

Acknowledgements

This project was funded by the Nuffield Foundation (EDU/40477 to GV and CN), but the views expressed are those of the authors and not necessarily those of the Foundation.

References

Andrews, M., & Vigliocco, G. (2010). The Hidden Markov Topic Model: A Probabilistic Model of Semantic Representation. *Topics in Cognitive Science*, *2*(1), 101–113. doi:10.1111/j.1756-8765.2009.01074.x

Audacity Team. Audacity® [Computer software]. Pittsburgh: Carnegie Mellon University. Retrieved from http://audacityteam.org/

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445–459.

Baron-Cohen, S., Golan, O., Wheelwright, S., Granader, Y., & Hill, J. (2010). Emotion word comprehension from 4 to 16 years old: a developmental survey. *Frontiers in Evolutionary Neuroscience*, *2*, 109. doi:10.3389/fnevo.2010.00109

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278. doi:10.1016/j.jml.2012.11.001

Begg, I., & Paivio, A. (1969). Concreteness and imagery in sentence meaning. *Journal of Verbal Learning and Verbal Behavior*, *8*(6), 821–827. doi:10.1016/S0022-5371(69)80049-6

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex (New York, N.Y. : 1991)*, *19*(12), 2767–96. doi:10.1093/cercor/bhp055

Bradley, M.M., & Lang, P.J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Brysbaert, M. (in press). Age of acquisition ratings score better on criterion validity than frequency trajectory or ratings ‘corrected’ for frequency. *Quarterly Journal of Experimental Psychology.*

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–11. doi:10.3758/s13428-013-0403-5

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*(6), 215–222. doi:10.1016/S1364-6613(00)01483-2

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505. doi:10.1080/14640748108400805

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale, 2nd edition*. Windsor, UK: NFER-Nelson.

Fox, J. R. (2004). A Signal Detection Analysis of Audio/Video Redundancy Effects in Television News Video. *Communication Research*, *31*(5), 524–536. doi:10.1177/0093650204267931

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23–64. doi:10.1038/sj.bdj.4803768

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211–44. doi:10.1037/0033-295X.114.2.211

Hamilton, A., Plunkett, K., & Schafer, G., (2000). Infant vocabulary development assessed with a British Communicative Development Inventory: Lower scores in the UK than the USA. *Journal of Child Language, 27*(3), 689-705.

James, C. T. (1974). The Role of Semantic Information in Lexical Decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 130-136. doi: 10.1037/0096-1523.1.2.130

Kiraly Jr., J. & Furlong, A. (1974). Teaching words to kindergarten children with picture, configuration, and initial sound cues in a prompting procedure. *The Journal of Educational Research, 67*(7), 295-298.

Kousta, S., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481. doi:10.1016/j.cognition.2009.06.007

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*(1), 14–34. doi:10.1037/a0021446

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*(3), 1065–81. http://doi.org/10.1037/a0035669

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–90. doi:10.3758/s13428-012-0210-4

Kuznetsova, A., Brockoff, P. B., & Christensen, R. H. B. (2016). Tests in Linear Mixed Effects Models. R package lmerTest version 2.0-30. Comprehensive R Archive Network (CRAN).

Landauer, T. K., & Dutnais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. doi: 10.1.1.184.4759

Li, Y., & Yu, D. (2015). Development of Emotion Word Comprehension in Chinese Children from 2 to 13 Years Old: Relationships with Valence and Empathy. *PloS One*, *10*(12), e0143712. doi:10.1371/journal.pone.0143712

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide*. Cambridge, UK: Cambridge University Press.

MacWhinney, B. (2000). *The CHILDES Project: Tools for anlyzing talk.*Mahwah, NJ: Lawrence Erlbaum Associates, 3rd edition.

McFalls, E. L., Schwanenflugel, P. J., & Stahl, S. A. (1996). Influence of word meaning on the acquisition of a reading vocabulary in second-grade children. *Reading and Writing, 8*(3), 235–250. http://doi.org/10.1007/BF00420277

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*(2), 408–19. http://doi.org/10.3758/s13423-014-0696-0

Newcombe, P. I., Campbell, C., Siakaluk, P. D., & Pexman, P. M. (2012). Effects of emotional and sensorimotor knowledge in semantic processing of concrete and abstract nouns. Frontiers in Human Neuroscience, 6, 275. http://doi.org/10.3389/fnhum.2012.00275

Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart and Winston.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience*, *19*(8), 1407–19. doi:10.1162/jocn.2007.19.8.1407

Psychology Software Tools, Inc. (2012). E-Prime 2.0 [Computer software]. Retrieved from http://www.pstnet.com.

R Core Team (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Richmond, M. G. & McNinch, G. (1977). Word learning: Concrete vs. abstract acquisition. *Perceptual and Motor Skills, 45*, 292–294

Rubin, D. C. (1980). 51 properties of 125 words: A unit analysis of verbal behavior. *Journal of Verbal Learning and Verbal Behavior*, *19*(6), 736–755. doi:10.1016/S0022-5371(80)90415-6

Schwanenflugel, P. J. (1991). Why are abstract concepts hard to understand? In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Shapiro, M. A. (1994). Signal Detection Measures of Recognition Memory. In A. Lang (Ed.), *Measuring Psychological Responses To Media Messages* (pp. 133–148). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Shulman, C., & Guberman, A. (2007). Acquisition of verb meaning through syntactic cues: A comparison of children with autism, children with specific language impairment (SLI) and children with typical language development (TLD). *Journal of Child Language*, *34*(02), 411. doi:10.1017/S0305000906007963

Siakaluk, P. D., Knol, N., & Pexman, P. M. (n.d.). Effects of emotional experience for abstract words in the Stroop task. *Cognitive Science, 38*(8), 1698–717. http://doi.org/10.1111/cogs.12137

Tomasello, M. (2010). *Origins of Human Communication*. MIT Press.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67*(6), 1176–90. http://doi.org/10.1080/17470218.2013.850521

Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: the role of emotion. *Cerebral Cortex (New York, N.Y. : 1991)*, *24*(7), 1767–77. doi:10.1093/cercor/bht025

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2014). Toward a theory of semantic representation. *Language and Cognition*, *1*(02), 219–247. doi:10.1515/LANGCOG.2009.011

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI) - Subject Baseline*. New York, NY: The Psychological Corporation: Harcourt Brace & Company. Retrieved from http://www.pathwaysstudy.pitt.edu/codebook/wasi-sb.html

Whaley, C. P. (1978). Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 143–154. doi:10.1016/S0022-5371(78)90110-X

Table 1

Lexical and sublexical characteristics of stimuli.

Variable	Concreteness category		t(23)	p
	Abstract	Concrete		
Concreteness[a]	328.6 (35.0)	530.1 (71.2)	-14.669	< 0.001
Length (n. of letters)	6.16 (0.9)	6.08 (0.8)	1.528	0.170
Valence[b]	4.89 (1.6)	4.97 (1.7)	-0.719	0.495
Age of Acquisition[c]	7.96 (2.4)	7.63 (2.4)	1.887	0.095
Log frequency[d]	9.33 (1.4)	9.16 (1.3)	1.216	0.264
CBBC frequency[e]	4.17 (0.74)	4.04 (0.76)	0.658	0.517
Familiarity[a]	511.04 (72.1)	510.63 (59.8)	0.03	0.977
Imageability[a]	400.96 (43.89)	531.29 (48.26)	-9.021	< 0.001
Arousal[b]	1.31 (0.93)	1.35 (0.99)	-0.315	0.762
ELP[d] – Accuracy	0.97 (0.05)	0.95 (0.08)	1.101	0.307
ELP[d] – RTs	649.65 (63.9)	659.50 (85.0)	-0.994	0.353

Note. Standard deviations appear in parentheses below means. a. Coltheart, 1981; b. Warriner et al., 2013; c. Kuperman et al., 2012; d. Balota et al., 2007; e. Van Heuven et al., 2014.

Table 2

Participants' age, gender and performance on standard tests of non-verbal reasoning and vocabulary.

Age Group	N	Male/ Female	Mean Age	Age range	Matrix Reasoning t-score	BPVS standard score
6-7 years	18	7/11	6.70 (0.51)	6.08-7.67	57.71 (9.50)	102.94 (9.06)
8-9 years	16	8/8	9.12 (0.60)	8.12-9.95	52.81 (8.19)	99.75 (6.01)
10-11 years	17	7/10	10.91 (0.57)	10.02-11.79	51.12 (5.58)	96.24 (11.82)

Note: Standard deviations appear in parentheses below means. Matrix reasoning t-scores have a normative mean of 50 and a standard deviation of 10. BPVS standard scores have a normative mean of 100 and a standard deviation of 15.

List of figure legends

Figure 1 – Percentage of abstract words (over total words) known at each age, from <4 to 12+ years. The histogram indicates the average percentage of abstract words known at different age groups. The dots indicate the percentage of abstract words known at age intervals = 0.1 years, from 2.7 to 14 years. The curve indicates the polynomial function used to calculate the inflection point (vertical line).

Figure 2 - Plot of the partial effect of valence (1 = negative; 5 = neutral; 9 = positive) as a predictor of concreteness (1 = abstract; 5 = concrete); data from regression analysis of 13,115 words. Grey area indicates 95% confidence intervals.

Figure 3 - Plot of the partial effect of valence (1 = negative; 5 = neutral; 9 = positive) on age of acquisition (in years) on 5,460 abstract words (left panel) and on 7,654 concrete words (right panel). Grey area indicates 95% confidence intervals.

Figure 4 - Trial timeline.

Figure 5 - Proportion of correct responses of 6-7 years old, 8-9 years old and 10-11 years old children at the lexical decision task as a function of words' concreteness and valence. Means and confidence intervals have been computed from a model including the significant predictors familiarity and CBBC frequency, and the three-way interaction between age group, concreteness and valence. Error bars represent 95% confidence intervals based on fixed-effects uncertainty.

Appendix A

Concrete and abstract words used in the lexical decision experiment

Abstract	Concrete
agility	beach
analogy	cancer
cause	caress
crisis	coast
custom	empire
danger	enemy
dozen	fatigue
dream	fencing
error	garden
forfeit	ground
goddess	leaflet
heaven	mineral
inquiry	missile
instinct	motor
length	movie
minute	pioneer
regency	range
revenge	resort
scheme	scholar
shame	scratch
style	spider
tragedy	square
trouble	sulphur
trust	waste

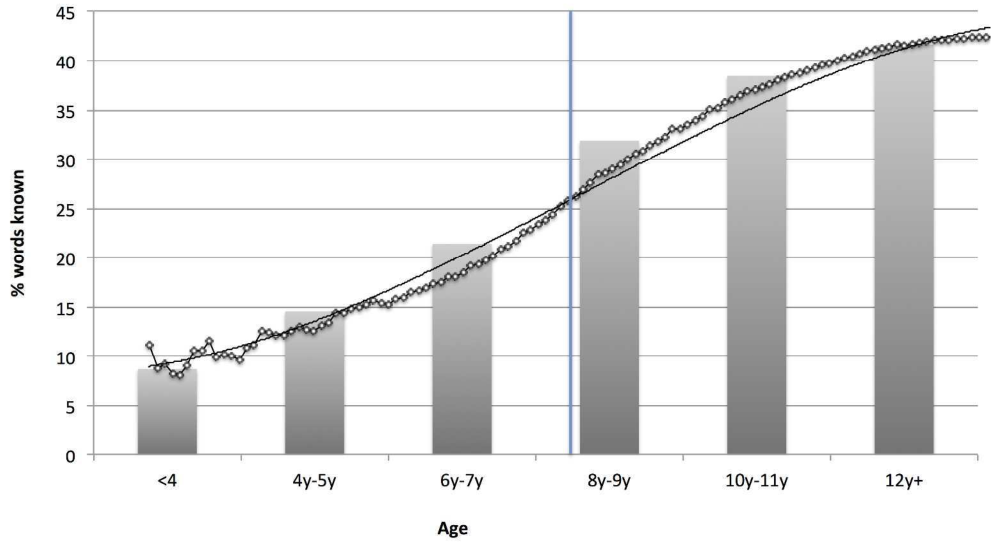

Figure 1 – Percentage of abstract words (over total words) known at each age, from <4 to 12+ years. The histogram indicates the average percentage of abstract words known at different age groups. The dots indicate the percentage of abstract words known at age intervals = 0.1 years, from 2.7 to 14 years. The curve indicates the polynomial function used to calculate the inflection point (vertical line).
Please insert Figure 1 about h
254x149mm (300 x 300 DPI)

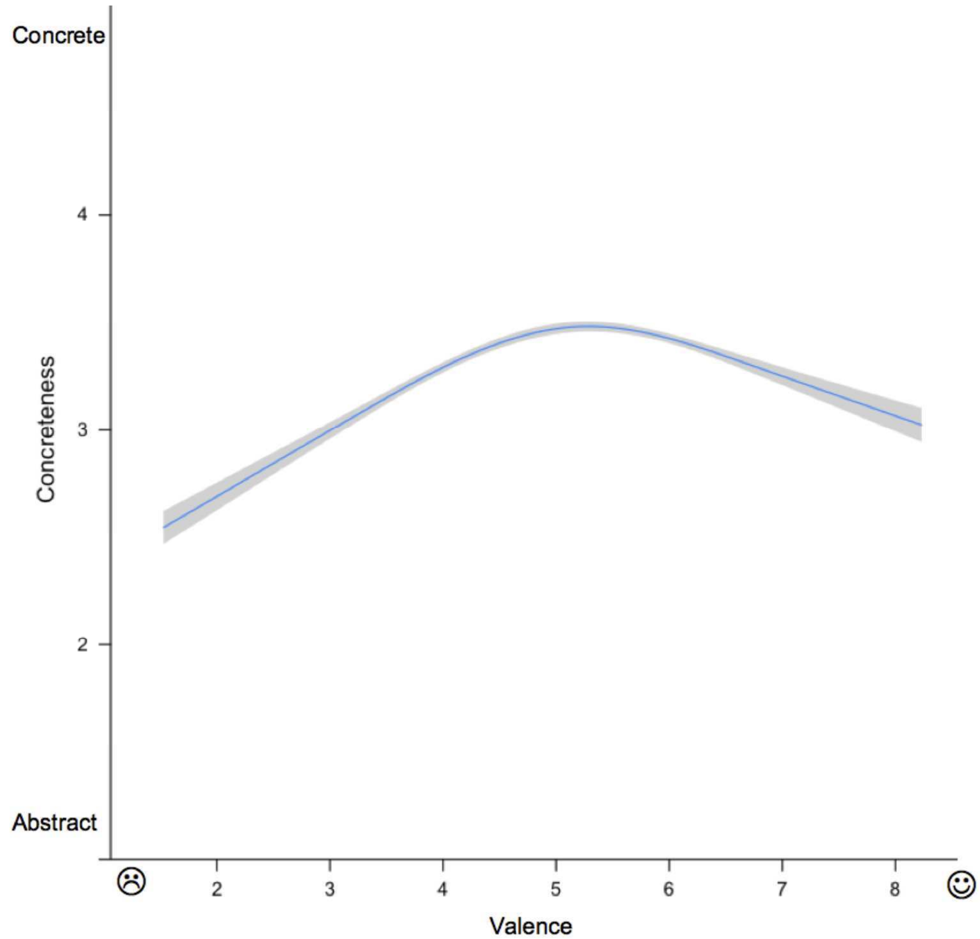

Figure 2 - Plot of the partial effect of valence (1 = negative; 5 = neutral; 9 = positive) as a predictor of concreteness (1 = abstract; 5 = concrete); data from regression analysis of 13,115 words. Grey area indicates 95% confidence intervals.
Please insert Figure 2 about h
150x142mm (300 x 300 DPI)

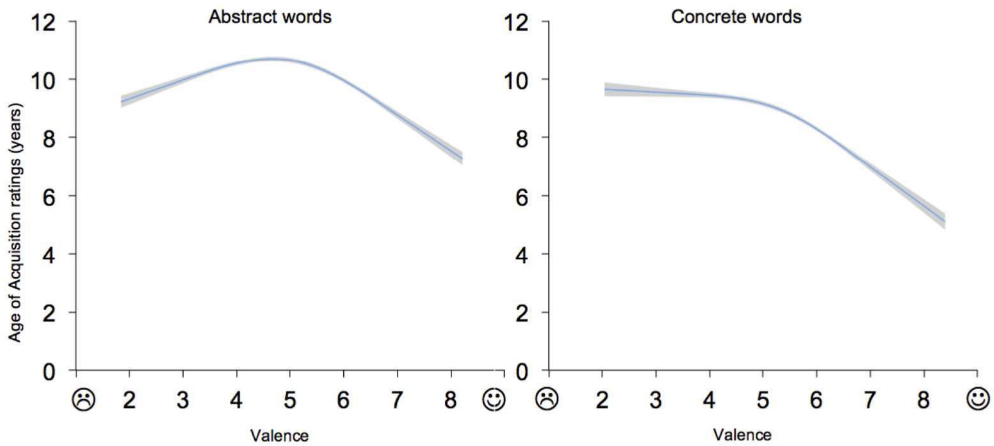

Figure 3 - Plot of the partial effect of valence (1 = negative; 5 = neutral; 9 = positive) on age of acquisition (in years) on 5,460 abstract words (left panel) and on 7,654 concrete words (right panel). Grey area indicates 95% confidence intervals.
Please insert Figure 3 about h
150x74mm (300 x 300 DPI)

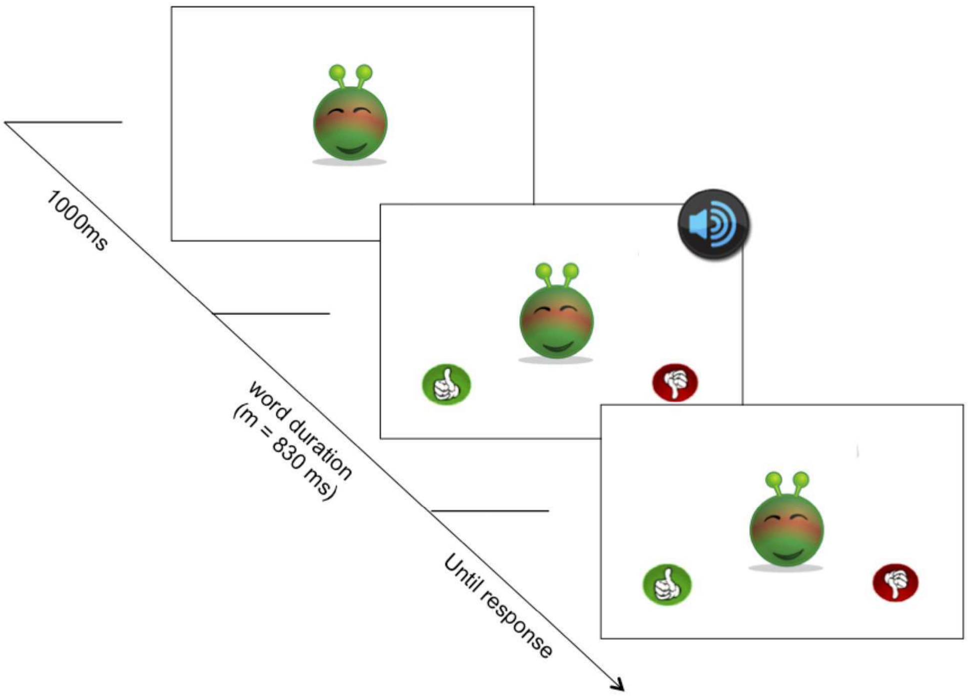

Figure 4 - Trial timeline.
Please insert Figure 4 about h
150x110mm (300 x 300 DPI)

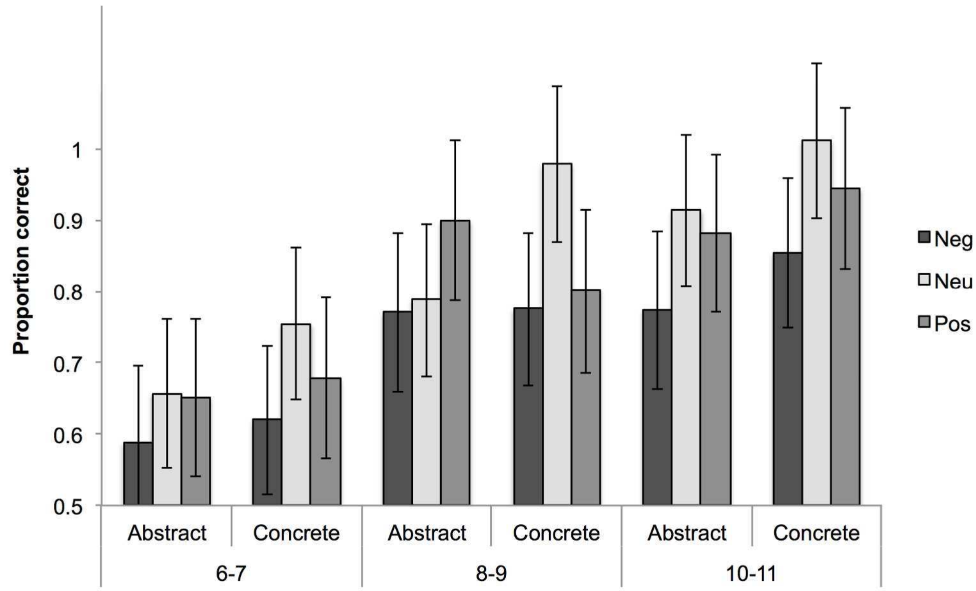

Figure 5 - Proportion of correct responses of 6-7 years old, 8-9 years old and 10-11 years old children at the lexical decision task as a function of words' concreteness and valence. Means and confidence intervals have been computed from a model including the significant predictors familiarity and CBBC frequency, and the three way interaction between age group, concreteness and valence. Error bars represent 95% confidence intervals based on fixed-effects uncertainty.
Please insert Figure 5 about h
197x120mm (300 x 300 DPI)