Age of acquisition estimates for 1,208 ambiguous and polysemous words

Maya M. Khanna · Michael J. Cortese

Published online: 16 November 2010
© Psychonomic Society, Inc. 2010

Abstract Age of acquisition (AoA) estimates are provided for 3,460 senses of 1,208 words (i.e., words with multiple meanings e.g., *duck*). The AoA rating estimates appear to be relatively consistent across participants. The Spearman-Brown split-half reliability coefficient is .95, while the correlations between each participant's ratings and the overall mean ratings yielded correlation coefficients between .325 to .794 with a mean of .69 ($SD = .10$). These estimates will be of use to those interested in: (a) the influence of AoA on word processing, (b) the influence of AoA on meaning access, (c) the structure of semantic memory, and (d) developmental trends in lexical ambiguity resolution. These AoA estimates can be downloaded from the Psychonomic Society's Web archive of norms, stimuli, and data at www.psychonomic.org/archive.

Keywords Age of acquisition · Ambiguity · Homophones · Polysemy

Introduction

In the past few years, a relatively large number of megastudies have been conducted to examine various

properties that influence word processing and memory for words (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Cortese & Fuggett, 2004; Cortese & Khanna, 2007, 2008). In these studies, participants respond to a large series of items (e.g., 3,000 monosyllabic words) by reading them aloud, rating them on a dimension (e.g., imageability) or making a lexical decision, among other tasks. In the megastudy approach, this large number of responses for a large number of stimuli allows the researcher to examine which properties of their corpus influence the participants' responses through the conduction of regression analyses. For example, in the studies reported by Balota et al. (2004) on monosyllabic words, it was found that a myriad of variables (e.g., length, objective frequency, subjective frequency, initial phoneme characteristics, and others) influenced participants' naming and lexical decision responses (Balota et al., 2004). Furthermore, another goal for megastudies is to provide access to the standard participant responses for the test items. This can be useful to other researchers who are trying to select stimuli that fit certain behavioral criteria.

In addition to this recent interest in megastudies, there also has been considerable interest in examining the influence of a word's age of acquisition (AoA, i.e., the age at which the meaning of a word is learned and understood by an individual) on the processing of the word. A recent study has provided a relatively comprehensive set of AoA ratings for most monosyllabic words in English (Cortese & Khanna, 2008). These ratings have been shown to be quite valuable in terms of predicting participant behavior in reading aloud and lexical decision measures of these words. Specifically, Cortese and Khanna (2007) found that as AoA increases, the naming latencies and lexical decision times also increase. Moreover, for both of these reaction time measures, AoA accounted for unique

Electronic supplementary material The online version of this article (doi:10.3758/s13428-010-0027-y) contains supplementary material, which is available to authorized users.

M. M. Khanna (✉)
Creighton University,
Omaha, NE, USA
e-mail: mayakhanna@creighton.edu

M. J. Cortese
University of Nebraska at Omaha,
Omaha, NE, USA

variance above and beyond 22 other factors (also see Brysbaert & Cortese, *in press*).

Prior to Cortese and Khanna's (2008) normative study on AoA, a number of other studies provided AoA estimates for words (e.g., Bird, Franklin, & Howard, 2001; Ferrand et al., 2008; Gilhooly & Logie, 1980; Stadthagen-Gonzalez & Davis, 2006). Interestingly, none of these studies took into account that there are many words in English that are lexically ambiguous (i.e., have multiple meanings, e.g., *trip*) or polysemous (i.e., have multiple meanings that are different, but related to one another, e.g., *break* can mean *to separate into pieces* or *to take a rest from work*). For some words, the multiple meanings were likely acquired at approximately the same age (e.g., *trip*), however, there are other words (e.g., *arm*) in which the multiple meanings likely were acquired at very different ages.

Ambiguous words have been used in a myriad of examinations of word recognition and language processing (see Simpson, 1994, for a review). One of the reasons that ambiguous words are so useful to researchers is because successful comprehension of these words requires the use and integration of context. Ambiguous words are certainly not an anomaly in language, and especially not in English. In fact, it easily could be argued that most words in English have more than one meaning or sense. That is, most words in English either have multiple distinct meanings (e.g., *bark* meaning *the outer covering of a tree* or *the sound that a dog makes*) or they have different, but interrelated meanings (e.g., *coffee breaks* and *glass breaks*; cf., Simpson & Burgess, 1988, also see Wordnet at <http://wordnet.princeton.edu/>). Thus, one could argue that it is important for future examinations of lexical-level properties to acknowledge the potential difference in processing of the multiple senses of ambiguous words, regardless of whether these senses represent distinct meanings or interrelated meanings of ambiguous words. To begin this endeavor, we have collected a set of ratings on the age of acquisition for 1,208 ambiguous word sets in which there are between two and six senses for each of these items. Thus, in total, we collected AoA ratings for 3,460 ambiguous word senses. We should emphasize that several previous studies have provided free association norms, dominance ratings, and measures of lexical access for the multiple senses of ambiguous words (cf., Gorfein & Weingartner, 2008; Twilley, Dixon, Taylor, & Clark, 1994; White & Abrams, 2004). However, these studies differ in the specific types of ambiguous words on which they focus. For example, White and Abrams (2004) collected free associations and dominance ratings for the multiple senses of homophones, which they defined as words that share the same phonology, but not the same orthography (e.g., *beach* and *beech*). While Gorfein and Weingartner (2008) also examined free associations to the written presentation of homophones, they, in addition, examined how participants

would spell these homophones when presented auditorily. Similarly to Gorfein and Weingartner (2008), we were interested in the properties of ambiguous word senses that could be examined in studies that used auditory and/or text presentation. Thus, we chose to examine both homographs and heterographic homophones. Homographs are words that share the same orthography. Sometimes these homographs share the same phonology (e.g., *bat* as a *flying mammal* or as *baseball equipment*), other times these homographs do not share the same phonology (e.g. *bass* as a *type of guitar* or as a *type of fish*). On the other hand, heterographic homophones share the same phonological code, but do not share the same orthography (e.g., *scent*, *sent*, and *cent*). We decided to focus on these ambiguous word types rather than focusing on the more narrow category of homonyms (i.e., words that share both orthography and phonology, but that have distinct meanings, e.g., *bat*) because most examinations of word processing are administered via text presentation or auditory presentation, but rarely in both modes. For example, a participant may be asked to make a lexical decision when presented with the text of *bass* or *bat*. This text-based presentation does not control the phonological code that the participant may assign to the letter string. Alternatively, a participant may hear /s nt/ in a free association task. In this case, the experimenter would not be able to control the orthographic code that the participant may assign to this sound.

We decided to collect subjective AoA ratings for the multiple senses of these ambiguous words for four reasons. First, in our previous examination (Cortese & Khanna, 2007), we concluded that AoA is a relatively robust semantic variable and so it seems reasonable to consider the AoA of the multiple senses independently. Second, several studies have suggested that access of the contextually appropriate meaning of an ambiguous word may depend on meaning frequency (i.e., dominance, Rayner & Frazier, 1989, for a review; however, see Martin, Vu, Kellas, & Metcalf, 1999, for a different view). Thus, because AoA estimates for words predict the speed of word processing (i.e., as measured via reading aloud and lexical decision tasks), it seems reasonable that the AoA of the individual senses of an ambiguous word will influence meaning access as well. If so, one could argue that these AoA norms for multiple ambiguous word senses have implications for broader theories of the architecture and operation of the linguistic system (e.g., Fodor, 1986). We address this point more fully in the Results and Discussion section.

A third reason for collecting these AoA estimates is that according to the semantic locus hypothesis or the developing semantic network hypothesis (Brysbaert, Van Wijnendaele, & De Deyne, 2000; and Steyvers & Tenenbaum, 2005,

respectively), early acquired concepts are stored in central semantic nodes and more-recently acquired concepts are established by activating and interconnecting those earlier-acquired semantic nodes. Because these early acquired concepts form the central semantic nodes from which other concepts branch, these early acquired concepts are activated more frequently than are later-acquired concepts. These theories are relevant and interesting for the study of ambiguous words because they relate to the debate about how the multiple senses of ambiguous words are represented in the lexicon. According to the shared hypothesis, the multiple senses share the same orthographic and/or phonological representations, but have separate semantic representations (cf. Cutting & Ferreira, 1999; Dell, 1990). On the other hand, in the independent representation hypothesis, the assumption is that the individual senses of the ambiguous words do not share any common representations and are stored in the lexicon in much the same way as would any two words that have distinct meanings, but overlapping orthographic and/or phonological features (cf., Caramazza, 1997; Harley, 1999). Thus, one could examine whether the AoA of the multiple senses of ambiguous words influences the processing of each other in order to indicate the representational overlap of the multiple senses of the ambiguous words.

A fourth reason we decided to collect AoA ratings for ambiguous word senses is because we felt that these ratings should be valuable for any researcher interested in ambiguity resolution, but especially useful for those interested in the development of lexical ambiguity resolution (e.g., Khanna & Boland, 2010). These ratings will allow the interested researcher to select ambiguous words for studies in which one or more meanings of the word are known by a desired age group of participants. Although AoA ratings are subjective and retrospective estimates, the high degree of inter-rater reliability of previous AoA ratings (both within and across studies) suggests that they can provide at least a relative measure of whether or not a word has been acquired by a group of children of a specific age (Cortese & Khanna, 2007).

Method

Participants

Participants were 60 undergraduate students from Creighton University ($n = 46$) and the University of Nebraska at Omaha ($n = 14$). There were 29 female participants and 31 male participants. The average age of our participants was 19.9 years. All participants were native English speakers. Participants received course credit in exchange for their participation.

Stimuli

We selected 3,460 ambiguous word senses and their meanings for inclusion in our stimuli set. These senses represented 1,208 word sets¹ with each set having an average of 2.86 senses included. We selected ambiguous words for inclusion in this study by examining several available databases. First, we used the University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 1998) for cues and their targets. In these norms, Nelson et al. (1998) reported a list of cue words and the associates that participants gave when presented with each of these cues. These cues included many ambiguous words, thus, many of these cues had lists of targets that indicated participants were accessing multiple meanings of the cue. For example, *nut* was a cue to which many of their participants indicated *crazy* as an associate and many other participants gave *squirrel* as an associate. Clearly, this indicates that *nut* has multiple senses that were accessed during their norming procedure. We went through the entire set of free association norms that were collected for 5,019 cue words and included in our ambiguous word list any cues that had associates given for two or more senses of the cue. We also included additional words that were included in a homograph meaning frequency study conducted by Twilley et al. (1994).

Furthermore, we included additional ambiguous words and senses that were not found in either the Nelson et al. (1998) or the Twilley et al. (1994) norms. While we were going through the Nelson et al. (1998) norms, we realized that there were several ambiguous words that served as cues but did not elicit associates to some of the subordinate senses of the words, but did elicit associations for us (likely because we were specifically trying to find multiple meanings for the cues). For example, *account* did not appear to elicit access to multiple meanings or senses in the Nelson et al. (1998) norms, but we do know that this word has multiple senses (e.g., a bank *account* and one's *account* or version of an event). Thus, we included these words in our list as well.

For each ambiguous word item on our list, we selected a word or brief definition to indicate to the participant which sense of the homograph he/she was to rate. Whenever possible, these brief definitions were based on the associates given in the Nelson et al. (1998) norms. For example, in their free association norms, the word *ace* produced the

¹ There was one word, *edge*, for which AoA ratings were collected, but for only one of its meanings (border). An alternative meaning to *edge* (to inch closer) was left out of the stimuli list, inadvertently. Despite our omission, we are including the rating information for this one meaning of *edge*. This item is assigned the 1209 ambiguous word set number.

associates, *cards*, *expert*, and *tennis* among other words. Thus, we used these associates (i.e., *cards*, *expert*, and *tennis serve*) as the basis for our three distinct definitions for *ace*. As described in detail, below, each rating trial included not only the presentation of a homograph (e.g., *ace*), but also the presentation of the associate or brief phrase that was to clarify which meaning of the ambiguous word we wanted the participant to rate.

We should note that the criteria we used to include items in our norming set resulted in a reasonably large and inclusive set of items. However, the dataset includes some items that others may not consider truly ambiguous or polysemous. For example, there are several items that are denominal verbs and deverbal nouns (e.g., *cage*, *table*, *catch*, etc.) in which both the noun and verb senses are included as separate items in the norming set. In addition, we include items in the norming set that may be considered metaphors, and not distinctly different meanings (e.g., *chill*, *cold*, etc.). We decided to include these items in the norming set because we felt that it might be useful to know the distinct AoA estimates for these different, although related, senses of these ambiguous words. For example, these normative values may be of interest to a researcher investigating the processing of these metaphoric or verb/noun sets of ambiguous word senses.

Furthermore, we included items in the dataset that were presented in both their monomorphemic (e.g., *arm*, *break*, and *admit*) and derived/inflected forms (e.g., *arms*, *breaks*, *admission*). Moreover, the inclusion of the monomorphemic as well as the inflected/derived forms could allow the investigation of the processes that underlie how and when these inflections/derivations are acquired as compared to the root forms of words. We should also note that others could argue that we did not include an exhaustive set of items that may be considered ambiguous. For example, our set does not include the words *mercury*, *atlas*, *colon*, *unlockable*, *undoable*, and others that may be considered polysemous. We did not overtly intend to omit these items, however, we did feel constrained to create a norming set that was not so extensive as to create an experiment that would overburden participants.

Due to the large number of stimuli (3,460) in this rating experiment, we had each participant rate one-half of the items. To do this, we created a master list that included all of the items. We then divided this list into four files; the assignment of items into these four files was determined through random assignment. Each participant rated the items included in two of these files, one per rating session. The file that each participant received in a session was determined via a random assignment and the trial presentation order within each session was also randomized.

Because each participant only rated half of the items, we had to run a complementary pair of participants in order to get a complete set of ratings for our stimuli. Thus, we obtained 30 ratings for each ambiguous word item from our 60 participants.

Equipment

Stimuli were presented on a 17-inch computer monitor that was controlled by a microcomputer. Stimulus presentation and data recording were coordinated via a program written in Turbo Pascal 7.0® (Borland International, Inc.) by the second author.

Procedure

We collected the ratings via microcomputers that were housed in research laboratories within the psychology department of each participant's university. Each participant engaged in two rating sessions that occurred on separate days, but that were less than 1 week apart. Each session lasted approximately 90 min and contained 865 trials. Each session consisted of three blocks of 288 or 289 trials. Participants were given the opportunity to take a break at the end of each trial block. On each trial, an ambiguous word was presented in the center of the screen. Immediately below each ambiguous word was a word or phrase that indicated the sense of the word to which we were referring. Underneath the word and its definition was a scale, ranging in value from 1 to 7. This scale appeared on every screen for easy reference. Participants entered their AoA ratings using the keypad on the right-hand side of the keyboard. The scale that we used was based on the one developed by Gilhooly and Logie (1980) and used again by Cortese and Khanna (2007, 2008). The instructions were modified to indicate that the participant should rate the AoA for the specific ambiguous word sense indicated on each trial. The instructions used in the current study appear below.

Instructions

This study is designed to determine the approximate age at which you learned different meanings of ambiguous words (i.e., words that have more than one meaning, e.g., *duck*). We acquire words throughout our lives. Some words are acquired at a very early age, some are acquired later, and others fall in between. The purpose of this study is to determine the approximate age for which 3,460 words have been acquired. On each trial of the study, you will be presented with a word along with one of its meanings. Try to estimate the age at which you acquired that word according to the scale shown at the

bottom of the screen. Please be sure that your rating is based only on the meaning of the word listed. On each trial, you will see a word along with one of its meanings. For example, you might see:

duck
to bend down

Try to estimate the age you learned this meaning for that word, according to the scale below. Word meanings acquired early should be given low ratings, and word meanings acquired later should be given a higher rating. For word meanings that you do not know, please assign a value of 7.

	1	2	3	4	5	6	7
AGE	0–2	2–4	4–6	6–8	8–10	10–12	13 +

When making your ratings, try to be accurate, but do not spend too much time on any one word. If you have any questions, ask the experimenter now. Otherwise, PRESS<ENTER>TO BEGIN THE STUDY.

On each trial screen, the participants were given one of the ambiguous word senses with a brief definition or associate listed under the word and the rating scale listed on the bottom of the trial screen. For example, when rating *duck* a participant would see:

	duck to bend down						
	1	2	3	4	5	6	7
AGE	0–2	2–4	4–6	6–8	8–10	10–12	13 +

Results and discussion

Upon obtaining our ratings, we created one rating data file for each complementary pair of participants. As indicated above, each participant rated half of the homograph items. Then, ratings were trimmed if the rating response time was very short. That is, we eliminated the ratings and their reaction times for any ratings that were made in less than 600 ms. We chose this reaction time cutoff in order to eliminate ratings that were made too quickly to warrant full consideration of the ambiguous word and the sense specified. We also chose this cutoff because we felt it still allowed relatively fast responses for easier items (e.g., *ball*, *a round object*). Choosing a larger cutoff value could potentially produce systematic bias in the estimates for these easy items. For example, if the cutoff value was raised to 1,000 ms, more responses to the easier items would be eliminated than to the more difficult items. In contrast, with the 600-ms cutoff value, it is likely that some premature responses were not removed, but these responses would be less likely to produce any systematic bias. We did not trim

the ratings due to overly long reaction times because our instructions did not emphasize making ratings as quickly as possible. From this screening procedure, we eliminated less than 1.6% of the ratings. We found that the mean AoA rating latency across homograph senses was 3,751.3 ms (SD = 1,054.7 ms). We note that this mean latency was 1,761.6 ms longer than the mean reported in the Cortese and Khanna (2008) AoA study in which words were rated without reference to any particular meaning.

After this trimming, we evaluated the degree to which each participant's ratings overlapped with those of the other participants, we initially calculated the average rating for each item. Then, we examined the correlation between each participant's ratings for each item rated and the average ratings from the other participants that rated that item. We found that there was a very high degree of inter-participant rating overlap, with a mean correlation coefficient of .69 (SD = .10). In addition, we examined the overall reliability of the AoA estimates by performing a Spearman-Brown split-half reliability analysis. We found that this reliability estimate was .95, which indicated a very high degree of similarity for the rating estimates across our participants. Finally, we also computed an average *z* score for each item. To calculate these item-level *z* scores, we computed each participant's mean AoA rating and the standard deviation of these ratings across items. We used these figures to calculate the *z* score for each item that the participant rated. We then calculated the average *z* score for each item across participants. This *z* score information provides another way to gauge the AoA for each item relative to the other items in our data set. We should note that the mean AoA ratings, reaction times, reaction time standard deviations, and *z* scores reported in the norm files do not include those ratings that were eliminated based on the above trimming procedures. Table 1 illustrates the correlations among our item measures (average AoA ratings, AoA rating reaction time, AoA rating reaction time standard deviation, and *z* scores for each item's AoA rating).

Description of ratings

The rating file that is available in the supplemental archive (www.psychonomic.org/archive) has seven distinct pieces of information for each ambiguous word sense. First, we list all of the 1,208 ambiguous word sets that were rated in our data set. Next, we list the individual senses of the words that were rated ($n = 3,460$). For each of these senses, we include the associate or phrases that indicated to the participants the ambiguous word sense that we wanted rated. Next, we include the average AoA ratings given for each ambiguous word sense. The average reaction time for the AoA rating for each sense is included as is the standard

Table 1 Item Level correlations between AoA rating estimates, average item rating reaction time, reaction time standard deviation, and item average z score

	1	2	3	4
1. Item Average AoA Rating Estimate	1	.048**	-.015	.995**
2. Average Item Rating Reaction Time	.048**	1	.815**	.052**
3. Reaction Time Standard Deviation	-.015	.815**	1	-.011
4. Item Average Z-score	.995**	.052**	-.011	1

**indicates $p < .01$

deviation of each reaction time. Finally, we include each item's AoA rating average z score.

Example analyses using ambiguous sense AoA estimates

The set of normative AoA estimates for the multiple senses of 1,208 ambiguous words should be useful for a variety of researchers interested in visual and auditory word recognition, lexical access, and lexical ambiguity resolution, among other topics. In terms of word processing, AoA has proven to be a relatively robust semantic variable that can account for reading aloud and lexical decision performance above and beyond a myriad of other semantic variables such as imageability, subjective frequency, and objective frequency (Cortese & Khanna, 2007). For example, predictions for word processing performance will be more accurate in light of the present AoA estimates that are sensitive to the particular senses of ambiguous words that are intended. In addition, this set of AoA estimates allows the examination of the relative influences of the various senses of ambiguous words on the processing of a particular ambiguous word.

For example, the naming latency and lexical decision times for a given ambiguous word may be influenced by the AoA of its multiple senses or these behavioral measures may be influenced by only the dominant sense of the ambiguous word. To illustrate this point and also to provide an example of how the current AoA estimates could be used, we chose to examine a subset of our ambiguous word senses in which relative frequency ratings were provided by Twilley et al. (1994). Specifically, we selected 314 sets of ambiguous word senses that were rated by our participants and those of Twilley and colleagues. We chose this set of norms, instead of other normative sets, because they contained the largest set of items that were also rated in our data set. In addition, these norms were for homographs that shared orthography and phonology and in which the senses were distinct (e.g., *bark* as a dog sound and the outer covering of a tree). Furthermore, the Twilley and colleague norms, which indicated the relative frequency of two or more meanings of homographs, allowed us to examine the influence on behavioral measures (e.g., lexical decision times, and naming latencies) of the AoA estimates for the dominant and subordinate senses of the homographs included in our data set. We then used the English Lexicon

Project (ELP; Balota et al., 2007) databases to gather normative data such as the speeded naming and lexical decision reaction times for these items. Please note that in the ELP participants were asked to read aloud or make a lexical decision via a text-based presentation of words with no indication of the meaning of the word. We are providing a file, also in the supplemental archive containing this subset of items along with the AoA estimates for the dominate and subordinate senses of the homographs, the ELP lexical decision time, and ELP naming reaction time.

To see if the AoA estimates for the dominant and subordinate senses, as indicated by the norms provided by Twilley and colleagues, each accounted for a proportion of the unique variance in these behavioral measures we conducted multiple regression analyses. In these analyses, the lexical decision reaction times and speeded naming reaction times served as two different outcome variables and the AoA estimates for the dominate and subordinate senses for the ambiguous word senses served as predictor variables. We found that a model including both the AoA estimates for the dominant and subordinate sense of the ambiguous senses accounted for 13.4% of the variance in lexical decision reaction times, adjusted $R^2 = .134$. In addition, both the dominant ($t = 5.50$, $p < 0.0001$) and subordinate senses' AoA estimates accounted for unique variance in lexical decision times ($t = 2.89$, $p = 0.004$). For naming latencies, we found that the combination of the dominant and subordinate AoA estimates accounted for 10.8% of the variance, again, with both the dominant and subordinate sense accounting for unique variance ($t = 3.32$, $p = 0.001$; and $t = 4.31$, $p < 0.0001$, respectively). Of course, these results should be interpreted with caution, as we did not attempt to control for other factors related to performance such as initial phoneme characteristics, length, frequency, etc. However, they do indicate that both the dominant and subordinate senses of these ambiguous words contribute to the processing of lexical forms. One could argue that this contribution from both senses indicates some degree of overlap in the representations of these senses in the lexicon.

Also, via the items from the Twilley et al. norms, we assessed the relationship between meaning dominance and AoA. We speculated that, in general, dominant meanings would be acquired before subordinate meanings. This hypothesis was confirmed via paired-samples t tests con-

ducted on mean AoA ratings for dominant and subordinate meanings. For the 314 homographs in the dataset, the mean AoA rating for the dominant meaning was 4.03 (SD = .87), and the mean AoA rating for the subordinate meaning was 4.73 (SD = .91), $t(313) = -11.49$, $p < 0.0001$, and for the z transformed means, $t(313) = -11.59$, $p < 0.0001$.

However, we also speculated that there may be interesting cases in which the subordinate meaning has been acquired prior to the dominant meaning. This characteristic was true for 80/314 items in the Twilley et al. norms. For instance, participants estimated that *tag – game* was acquired at an earlier age (mean AoA rating = 2.75) than *tag – label* (mean AoA rating = 4.32). If it turns out that AoA relates to the access of different meanings/senses, these items would be of particular interest. For example, researchers interested in the role of context on lexical ambiguity resolution could use AoA ratings for the multiple senses of ambiguous words to further address the subordinate bias effect (SBE, see e.g., Duffy, Morris, & Rayner, 1988; Martin et al., 1999). The SBE occurs when reading times are longer for ambiguous words in which context indicates the subordinate sense is intended as compared to when the dominant sense is supported. One might hypothesize that context may serve to reduce or eliminate the effect more easily when the subordinate meaning has been acquired earlier rather than later. This would be especially important because it would suggest a more limited role of context than previously thought. However, if context was found to eliminate the SBE even for those items whose subordinate meaning was acquired at a later age, this would suggest a prominent role for context in ambiguity resolution. In other words, as stated in the Introduction, these norms may provide the ability to test broader theories that relate to the modularity/interaction debate.

Thus, we hope it is clear that these AoA estimates may be considered useful for many different types of researchers. These estimates not only allow researchers to make better predictions about word processing, but they may also lead to a better understanding of how the multiple senses of ambiguous words influences one another. Furthermore, to an ambiguity researcher, an understanding of how the AoA ratings of the multiple senses of ambiguous words influences the processing of words can inform models that describe how ambiguous words are represented in the mental lexicon (cf. Miozzo, Jacobs, & Singer, 2004) and may provide insights into the operations of lexical processing, and linguistic architecture, in general.

Acknowledgements We would like to thank Kristin Daniels, Brendan Murphy, and Elizabeth Beverlin for their assistance in collecting the data. We would also like to thank Roberto G. de Almeida and an anonymous reviewer for their helpful comments on earlier versions of this article.

References

- 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. doi:10.1037/0096-3445.133.2.283
- 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.
- Bird, H., Franklin, D., & Howard, S. (2001). Age of acquisition and imageability ratings for a large set of words, including verb and function words. *Behavior Research Methods, Instruments, & Computers*, 33, 73–79.
- Brysbaert, M. & Cortese, M.J. (in press). Do the effects of familiarity and age of acquisition survive better frequency norms? *Quarterly Journal of Experimental Psychology*.
- Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, 104, 215–226. doi:10.1016/S0001-6918(00)00021-4
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208. doi:10.1080/026432997381664
- Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3, 000 monosyllabic words. *Behavior Methods and Research, Instrumentation, & Computers*, 36, 384–387.
- Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2, 342 words. *The Quarterly Journal of Experimental Psychology*, 60, 1072–1082.
- Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition rating for 3, 000 monosyllabic words. *Behavior Research Methods*, 40, 791–794. doi:10.3758/BRM.40.3.791
- Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 318–344. doi:10.1037/0278-7393.25.2.318
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313–349. doi:10.1080/01690969008407066
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27, 429–446. doi:10.1016/0749-596X(88)90066-6
- Ferrand, L., Bonin, P., Meot, A., Augustinova, M., New, B., Pallier, C., & Brysbaert, M. (2008). Age-of-acquisition and subjective frequency estimates for all generally known monosyllabic French words and their relation with other psycholinguistic variables. *Behavior Research Methods*, 40, 1049–1054. doi:10.3758/BRM.40.4.1049
- Fodor, J. (1986). The modularity of mind. In Z. W. Pylyshyn & W. Demopoulos (Eds.), *Meaning and cognitive structures: Issues in the computational theory of mind* (pp. 3–18). Westport, CT: Ablex Publishing.
- Gilhooly, K. J., & Logie, R. H. (1980). Meaning-dependent ratings of imagery, age of acquisition, familiarity, and concreteness for 387 ambiguous words. *Behavior Research Methods & Instrumentation*, 12, 428–450.
- Gorfein, D. S., & Weingartner, K. M. (2008). On the norming of homophones. *Behavior Research Methods*, 40, 522–530. doi:10.3758/BRM.40.2.522
- Harley, T. A. (1999). Will one stage and no feedback suffice in lexicalization? *The Behavioral and Brain Sciences*, 22, 45.
- Khanna, M. M., & Boland, J. E. (2010). Children's use of language context in lexical ambiguity resolution. *The Quarterly Journal of Experimental Psychology*, 63, 160–193. doi:10.1080/17470210902866664

- Martin, C., Vu, H., Kellas, G., & Metcalfe, K. (1999). Strength of discourse context as a determinant of the subordinate bias effect. *The Quarterly Journal of Experimental Psychology*, 52, 813–839. doi:[10.1080/027249899390819](https://doi.org/10.1080/027249899390819)
- Miozzo, M., Jacobs, M. L., & Singer, N. J. W. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology*, 21, 840–866. doi:[10.1080/02643290342000573](https://doi.org/10.1080/02643290342000573)
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. <http://www.usf.edu/FreeAssociation/>
- Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 779–790. doi:[10.1037/0278-7393.15.5.779](https://doi.org/10.1037/0278-7393.15.5.779)
- Simpson, G. B. (1994). Context and the processing of ambiguous words. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 359–374). San Diego, CA: Academic Press.
- Simpson, G. B., & Burgess, C. (1988). Implications of lexical ambiguity resolution for word recognition and comprehension. In S. I. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology & artificial intelligence* (pp. 271–288). San Mateo, CA: Morgan Kaufmann Publishers, Inc.
- Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability and familiarity. *Behavior Research Methods*, 38, 598–605.
- Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29, 41–78. doi:[10.1207/s15516709cog2901_3](https://doi.org/10.1207/s15516709cog2901_3)
- Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms for relative meaning frequency for 566 homographs. *Memory & Cognition*, 22, 111–126.
- White, K. K., & Abrams, L. (2004). Free association and dominance ratings of homophones for young and older adults. *Behavior Research Methods, Instruments, & Computers*, 36, 408–420.