University of Nebraska at Omaha
DigitalCommons@UNO

Psychology Faculty Publications

Department of Psychology

6-22-2020

AoA effects in reading aloud and lexical decision: Locating the (semantic) locus in terms of the number of backward semantic associations

Michael J. Cortese

Sean Toppi

Jonathan Santo

Follow this and additional works at: https://digitalcommons.unomaha.edu/psychfacpub

Part of the Psychology Commons

Please take our feedback survey at: https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE

AoA effects in reading aloud and lexical decision: Locating the (semantic) locus in terms of the number of backward semantic associations

Michael J Cortese[1], Sean Toppi[1], Maya M Khanna[2] and Jonathan B Santo[1]

[1]University of Nebraska Omaha, Omaha, NE, USA

[2]Creighton University, Omaha, NE, USA

Corresponding author:

Michael J Cortese, Department of Psychology, University of Nebraska Omaha, 6001 Dodge Street, Omaha, NE 68182, USA. Email: mcortese@unomaha.edu

Abstract

In the present study, we analyse data from the English Lexicon Project to assess the extent to which age of acquisition (AoA) effects on word processing stem from the number of semantic associations tied to a word. We show that the backward number of associates (bNoA; that is, the log transformed number of words that produce the target word in free association) is an important predictor of both lexical decision and reading aloud performance, and reduces the typical AoA effect as represented by subject ratings in both tasks. Although the AoA effect is reduced, it remains a significant predictor of performance above and beyond bNoA. We conclude that the semantic locus of AoA effects can be found in the number of backward connections to the word, and that the independent AoA effect is due to network plasticity. We discuss how computational models currently explain AoA effects, and how bNoA may affect their processing.

Keywords

Age of acquisition; number of backward associations; semantic locus hypothesis

Age of acquisition (AoA) is a general and strong predictor of word processing performance in tasks such as reading aloud and lexical decision, among others (see, for example, Cortese & Schock, 2013; Juhasz, 2005). AoA refers to the estimated age at which people typically learn a word. AoA is usually considered on a relative scale. For example, *top* is a word that is often acquired at a relatively early age as compared with a word such as *apex* which is typically acquired much later. Because AoA influences performance on measures of word processing, it is important to determine the loci of AoA effects because it has important implications for theories of semantic memory, lexical organisation, and word processing more generally.

One issue with the representation of AoA concerns how best to estimate it. Thus far, subjective ratings have been the most common way to measure AoA. That is, adults estimate the age at which they learned each of a series of words, some that were acquired many years ago. Participants are asked to perform these subjective ratings, often times, for hundreds if not thousands of words. In fact, Kuperman et al. (2012) collected such ratings via Amazon Mechanical Turk for 30,000 English words across 1,906 participants. Of course, the subjective nature of this method leaves open the possibility that the ratings are influenced by other lexical variables. While previous research has shown fairly strong correlations between the subjective AoA and more objective measures of AoA (see, for example, Carroll & White, 1973), the correlations are not perfect, and AoA correlates with many other lexical variables such as frequency, imageability, and orthographic length (Cortese & Khanna, 2008). Thus, if people are influenced by these other variables when making their estimates of AoA, it is possible that some proportion of AoA effects on word processing measures is actually due to these or other variables.

Consider word frequency. Words rated as earlier acquired tend to be more frequently encountered in the language. For example, the early acquired *top* is quite frequent as compared with the much later acquired *apex*. Accordingly, Zevin and Seidenberg (2004) proposed that frequency trajectory might be a better way to conceptualise AoA than subjective ratings. Frequency trajectory refers to the pattern of frequency associated with words across development. Some words (e.g., *potty*) are more frequently encountered early in development, some words are more frequently encountered later in development (e.g., *merlot*), and other words maintain a fairly constant frequency trajectory (e.g., *spoon*). And so, a word may be acquired at an early age when it is encountered more frequently early during development than when it is encountered less frequently early on. However, frequency trajectory has not been a strong predictor of word processing performance, and it has little influence on AoA's relationship with word processing (Brysbaert, 2017; Cortese & Khanna, 2007). Moreover, AoA effects are robust after more recent and accurate frequency norms are used to estimate frequency (see Brysbaert & Cortese, 2011). In other words, one cannot simply reduce AoA to either frequency or frequency trajectory.

Instead, we might better understand the nature of AoA if we look at semantic variables that relate to AoA. The present studies were motivated by the hypothesis that AoA has a semantic basis (e.g., Brysbaert et al., 2000; Steyvers & Tenenbaum, 2005; Van Loon-Vervoon, 1989). According to Steyvers and Tenenbaum's small world network theory, for example, new concepts are acquired incrementally via establishing associations with the meanings of previously acquired words. For example, when people encounter *apex*, they will likely make the association between it and *top*. Thus, the earlier acquired word (e.g., *top*) is processed at the semantic level when a later acquired word (e.g., *apex*) that shares semantic information is encountered. Because picture naming and lexical decision are assumed to rely on semantic information more than reading aloud, the semantic locus hypothesis can explain the differential effect of AoA across these tasks (Juhasz, 2005). More specifically, Steyvers and Tenenbaum (2005) propose that when a word is acquired, a node representing the concept of the word is established, and connections from it to a subset of nodes in a semantic network also form. Over the course of lexical development, certain early acquired words become hubs, and novel representations then connect to these earlier acquired hubs.

If new concepts are formed by making associations with already established concepts, then one should find that earlier acquired words (e.g., *top*) are produced more often in free association tasks (e.g., Nelson et al., 1998) than later acquired words (e.g., *apex*). Indeed, 65 other words generate *top* in free association, whereas only 3 other words produce *apex*. More generally, in a study of semantic associations of 1,117 Dutch words, De Deyne and Storms (2008) reported a correlation of –.61 between the AoA of target words and the (log transformed) back- ward number of associates (bNoA; that is, the words that led to the production of the target words in free association) of that target word. In contrast, AoA was not related to the number of forward associates (i.e., words produced by the target word in free association; $r = .03$). In other words, later acquired words produce just as many associates in free association as earlier words. In analyses of 3,055 English words, Schock, Cortese, Khanna, and Toppi (2012) replicated this pattern to a remarkable degree (the correlations were –.62 and .03, respectively), even though De Deyne and Storms (2008) utilised a continuous association task where multiple associates to a target are generated, whereas the Nelson et al. (1998) norms are based on a single response association task.

Clearly, the semantic locus hypothesis makes the pre- diction that at least part of the AoA effect can be reduced to the number of backward associates of the target word. Of course, there are other theories that offer viable explanations of AoA effects, but we save our discussion of these theories for the general discussion. In the present study, we examine how AoA relates to the number and direction of semantic associations of a word and whether these associations account for some AoA-related variance in reading aloud and lexical decision performance for 2,940 words. We extracted *z*-score transformed reaction times (i.e., *z*RTs) and accuracy measures from the English Lexicon Project (ELP). A semantic association occurs between two words when one word (e.g., *apex*) generates the other word (e.g., *top*) in free association (Nelson et al., 1998). The direction of the association in the case of *apex* and *top* is one way. Specifically, *apex* generates *top* in free association, but *top* does not generate *apex*. In terms of the terminology that we use throughout this article, *top* is a forward associate of *apex*, and *apex* is a backward associate of *top*. We note, that, in this case, *apex* is a later acquired word and *top* is an earlier acquired word.

Interestingly, to the best of our knowledge, no one has simultaneously explored the relationship among AoA, the number of forward and backward (or reciprocal; that is, words that are associated in both the forward and back- ward directions) associates of a target word, and word processing performance. We note that other studies (e.g., Balota et al., 2004; Steyvers & Tenenbaum, 2005; Yap & Balota, 2009; Yap et al., 2011) have included semantic neighbourhood or number of semantic associates as predictors, but those studies have not differentiated the direction of the associations in the way we have done in the present study. For example, Balota et al. (2004) assessed connectivity as the sum of the number of associates generated by a target (i.e., the forward associates) and the number of other words that generate the target (i.e., the backward associates). In the Nelson et al. (1998) norms, we find that *top* generates 11 associates, and, as mentioned, 65 other words generate *top*. Therefore, the connectivity value utilised by Balota et al. would be (the log of) 76. Balota et al. found that, in a sample of 1,625 monosyllabic words, connectivity is related to lexical decision reaction time (RT) but not reading aloud RT. In the present study, we parsed out this connectivity value by separating the influences of forward and backward associations. We did this because, as we indicate above, AoA is not related to the number of forward associates of a word, but it is related to the number of backward associates. Furthermore, we hypothesise that forward associates and backward associates likely have different influences on models of word processing. In addition, we thought it was important to acknowledge that some pairs of words act as both forward and backward associates of one another. For example, *top* is a forward associate of *hat*, and *hat* is a forward associate of *top*; we refer to these as reciprocal associates. Therefore, we also examine the influence of reciprocal associates on word processing measures.

Specifically, we examine the influence on word processing from three distinct measures of the number of associations a word has. First, we examine bNoA. This variable refers to the number of words that generate a tar- get word and is moderately associated with AoA (De Deyne & Storms, 2008; Schock, Cortese, & Khanna, 2012). Due to the log-linear relationship between AoA and bNoA, we log transformed bNoA for the purposes of the present study. Second, we examine the influence of for- ward number of associates (i.e., fNoA). This variable refers to the number of words that a target word itself generates in free association and is not strongly associated with AoA (De Deyne & Storms, 2008; Schock, Cortese, Khanna, & Toppi, 2012). For example, *top* generates 11 other words in free association, and *apex* generates 16 other words. Third, we examine the reciprocal number of associates (i.e., rNoA). This variable refers to the common forward and backward associates of a word. For example, *top* has six reciprocal associates (*bottom, hat, above, spin, cover*, and *shirt*), and *apex* has one reciprocal associate (*summit*). AoA is also correlated with rNoA, but the correlation is a bit weaker than that between AoA and bNoA (see below). Because bNoA and rNoA are highly correlated, we analysed these variables separately.

We felt that the examination of rNoA was needed because it is not clear how computational models of reading aloud would be influenced by bNoA. Most computational models do not assume much if any semantic influence in the reading aloud process. This is simply because, to read aloud a word, one needs to translate an orthographic code into a phono- logical code that is articulated. The influence of semantics tends to be minimal in this task (Cortese & Khanna, 2007). In fact, neither the dual route cascade (DRC; Coltheart et al., 2001) model of reading aloud nor the connectionist dual process (CDP++; Perry et al., 2010) model implement any semantic system at all. A recent exception that implements a semantic system into its computational framework comes from the bilingual word recognition model Multilink, pro- posed by Dijkstra et al. (2019). Within Multilink, there is a semantic system that interacts with both orthographic and phonological representations, and thus, the model assumes that reading aloud of a word may be influenced by its reciprocal semantic associates. On the contrary, backward associates would not be expected to influence reading aloud performance in Multilink unless a forward association was also present. For example, when *top* is read, the lexical (i.e., orthographic, phonological) representations for it would be activated, and this activation would then spread to the semantic system, activating *top*'s forward associates. As we noted above, later acquired words have just as many for- ward associates as earlier acquired words. However, importantly, the orthographic and phonological lexical representations for *top* would receive top-down activation from the semantic system via the reciprocal associates. In this case, *bottom, hat, above*, and so on would be activated from the bottom-up (via activation from *top*) and would reinforce activation of *top* either through top-down activation from the semantic level to the orthographic and phono- logical levels or via *top*'s semantic representation. In contrast, the later acquired word *apex* would benefit in a similar manner for only the single reciprocal associate *summit*. Like this example, earlier acquired words do tend to have more reciprocal associates than later acquired words. In the current study of 2,940 words, $r = –.485$ between log 10 rNoA and AoA. We note that the correlation is stronger between log 10 bNoA and AoA ($r = –.614$ in the current study) than log 10 rNoA and AoA.

Log 10 rNoA and log 10 bNoA are strongly correlated ($r = .812$).

We also note that the activation process described above for Multilink would be similar for parallel-distributed-processing (PDP) models (e.g., Chang et al., 2019; Harm & Seidenberg, 2004; Plaut et al., 1996), but representations would be distributed rather than localised. In particular, we think that the model proposed by Chang et al. (2019) rep- resents an important extension to the PDP framework because it incorporates a semantic network in addition to orthographic and phonological networks. This addition is important because it allows for the simultaneous computation of a phonological code and a semantic code from an orthographic input. Furthermore, through interactivity, the computation of a code at any one level could be influenced by activity occurring at any of the other levels, although Chang et al. do not implement interactivity from semantics or phonology towards orthography. Thus, in the model, the computation of a semantic code during reading occurs directly via orthography and also indirectly via an orthog- raphy-to-phonology-to semantic pathway. Furthermore, through interactive activation, the computation of a semantic code is influenced by activity accruing in the phono- logical network, and the computation of a phonological code is influenced by activity accruing in the semantic net- work. Interactivity between semantic and/or phonological levels and the orthographic level was not implemented, but theoretically, one would assume interactivity among all representational levels. In general, we think that the PDP approach provides a very good theoretical framework for explaining AoA effects (see Cortese & Khanna, 2007; Cortese & Schock, 2013) on measures of word processing. We discuss this approach in more detail in the "Results and discussion" section.

In the current study, we accessed lexical decision and reading aloud RT (*z*RT) and accuracy measures from the ELP for 2,940 monosyllabic and disyllabic words. The words that we selected for inclusion in this study were chosen because they had one or two syllables; had at least one forward associate, one backward associate, and one reciprocal associate; and because predictor values for standard predictors (e.g., frequency, AoA, imageability) were available. To assess the influence of AoA and the NoA variables (fNoA, bNoA, and rNoA) on performance, we applied stepwise multiple regression. In our analyses, initial phoneme was entered in the first step, and sublexical and lexical factors and imageability were entered at the second step. At the third step, the NoA variables were entered, and AoA was assessed in the final step.

We predicted that because lexical decision is thought to rely substantially on semantic information (Cortese & Khanna, 2007), all NoA variables would be expected to affect performance on that task. When a word is encountered, it will activate the representations of its associates from the bottom-up. Specifically, in dual route models, the orthographic and phonological representations will be activated by the printed word and will serve as input to the semantic system which will activate associates to the target word. If the target word activates a larger number of associates, this information would be useful in discriminating words from nonwords as nonwords do not have semantic associates. We assume that this process would be similar for distributed models (e.g., Plaut et al., 1996) but would involve distributed representations at the ortho- graphic, phonological, and semantic levels. In addition, rNoA could influence performance due to feedback from semantics to orthographic and phonological systems which would result in increased activity at the lexical level for words having relatively more rNoAs. It is not entirely clear how bNoA would affect performance in these models, except via rNoA.

For reading aloud, one might expect the effects of the NoA variables to be somewhat weaker than lexical decision because there is less emphasis on semantic information when reading aloud. Thus, the main way in which these NoA variables would influence contemporary models of reading aloud would be through recurrent activity from the semantic system to orthographic and phonological representations for words with relatively more rNoAs.

A central issue involves the degree to which the semantic locus of the AoA effect can be reduced to bNoA. If it can, then we would expect a substantial reduction of the AoA effect when bNoA is controlled versus when it is not. So, to address this issue, additional analyses were con- ducted without bNoA to compare the AoA effect with and without bNoA in the regression model. Regardless, if AoA remains significant after controlling for bNoA, then the time at which a word is learned is still an important characteristic that affects word processing.

Method

Materials

Stimulus characteristics for the 2,940 words in the studies can be found in Table 1, and the correlation matrix for the Steps 2 to 4 predictor variables can be found in Table 2.

Table 1. Stimulus characteristics for the 2,940 words utilised in the analyses.

Characteristics	*M*	*SD*
AoA	3.93	0.89
fNoA	14.12	5.15
bNoA	16.95	25.09
Log bNoA	0.95	0.49
rNoA	4.07	2.80
Log rNoA	0.51	0.31
Number of syllables	1.44	0.50
Length	5.11	1.33
Subtitle frequency	145.07	972.38
Log subtitle frequency	2.94	0.78
LOD	1.81	0.51
Imageability	4.91	1.36

AoA: age of acquisition, fNoA: forward number of associates, bNoA: backward number of associates; rNoA: reciprocal number of associates; LOD: Levenshtein orthographic distance.
Subtitle frequency refers to the Brysbaert and New (2009) per million values. In each case, log refers to the log base 10 transformation of the base value.

Table 2. Correlation matrix of the predictor variables used in the analyses.

	AoA	fNoA	Log bNoA	Log rNoA	Length	Log frequency	LOD	Imageability	Number of syllables
AoA	1.0	.03	–.614**	–.485**	.387**	–.582**	.382**	.226**	.404**
fNoA		1.0	.161**	.278**	.089**	.129**	.039*	–.102**	.021
Log bNoA			1.0	.812**	–.219**	.672**	–.245**	.070**	–.200**
Log rNoA				1.0	–.165**	.467**	–.186**	.145**	–.165**
Length					1.0	–.212**	.828**	.072**	.690**
Log frequency						1.0	–.219**	.226**	–.168**
LOD							1.0	.068**	.662**
Imageability								1.0	.118**

AoA: age of acquisition, fNoA: forward number of associates, bNoA: backward number of associates; rNoA: reciprocal number of associates; LOD: Levenshtein orthographic distance.
* $p < .05$
** $p < .01$

AoA scale

AoA ratings were taken from Cortese and Khanna's (2008) monosyllabic database and Schock et al.'s (2012) disyllabic database. In both studies, participants estimated the age at which they acquired each of 3,000 words on a 1 to 7 scale, where 1 indicated some time between 0 and 2 years of age, and each successive number on the scale incremented the AoA by 2 years until the rating of 7 which rep- resented the age of 13 or older.

NoA variables

All of the NoA variables were derived from the Nelson et al. (1998) norms. Due to their positive skew, we log transformed bNoA and rNoA for the analyses. fNoA did not have this problem and so was analysed without any transformation.

Control variables

The control variables were entered in the first two steps. Initial phoneme, in terms of 13 features (e.g., voicing, stop, bilabial), was dichotomously coded and entered in the first step. Log subtitle frequency (Brysbaert & New, 2009), number of syllables, orthographic length,[1] and imageability (Cortese & Fugett, 2004; Schock, Cortese, & Khanna, 2012) were entered in the second step.

Key predictors

AoA (Cortese & Khanna, 2008; Schock et al., 2012) and the NoA variables were entered in the third step. A high correlation between log bNoA and log rNoA ($r = .812$, $p < .001$) prevented us from assessing these variables together. Therefore, bNoA and rNoA were analysed separately. Fortunately, the correlations between fNoA and the other NoA variables were not problematic in terms of col- linearity, so fNoA was included in every analysis at the third step. As previously mentioned, to assess the reduction of the AoA effect in terms of bNoA, separate analyses were conducted with and without bNoA in the model.

Results and discussion

Six regression analyses (three for each task) were per- formed with RT as the dependent variable (DV; see Tables 3 and 4), and six analyses were performed with accuracy as a DV. We report only the results on RT for three reasons. First, the theoretical interest is mainly based on the RT data. Second, the pattern of all the critical variables (fNoA, log bNoA, log rNoA, and AoA) reported for the RT analyses is not qualified in any way by the accuracy results. Third, the accuracy results show the same, albeit smaller in magnitude, patterns as the RT results for all of the critical variables. The data used for all of the analyses are provided in Open Science Framework (OSF; https://osf.io/z8yq6/).

Table 3. Results from the regression analyses conducted on reading aloud zRTs.

Step 1			
Initial phoneme			
R^2	.274**		
Step 2			
Log frequency	−.306**		
Length	.258**		
Number of syllables	.060**		
Image	−.081**		
R^2	.485**		
Step 3	With bNoA	With rNoA	With neither bNoA nor rNoA
fNoA	−.015	−.008	−.029*
bNoA	−.140**		
rNoA		−.082**	
R^2	.495**	.490**	.486**
AoA	.134**	.150**	.164**
R^2	.502**	.499**	.496**
ΔR^2	.007**	.009**	.01**

zRTs: z-score transformed reaction times; fNoA: forward number of associates, bNoA: backward number of associates; rNoA: reciprocal number of associates; AoA: age of acquisition.
* $p < .05$
** $p < .01$

The results of the present studies indicate that (1) bNoA significantly relates to RT even after controlling for many factors related to performance; (2) the effect of the NoA variables is larger in lexical decision than reading aloud; (3) the AoA effect is reduced when bNoA is controlled; (4) the reduction in AoA, due to bNoA, is larger in lexical decision than reading aloud; and (5) the AoA effect remains a significant predictor after controlling for bNoA. Finally, we note that due to the high correlation between bNoA and rNoA, we were unable to fully distinguish the two variables. However, in all cases, bNoA was more strongly related to performance than rNoA, and the AoA effect was always reduced more when bNoA was controlled than when rNoA was controlled.

Table 4. Results from the regression analyses conducted on lexical decision zRTs.

Step 1			
Initial phoneme			
R^2	.008**		
Step 2			
Log frequency	−.588**		
Length	.077**		
Number of syllables	.132**		
Image	−.161**		
R^2	.405**		
Step 3	With bNoA	With rNoA	With neither bNoA nor rNoA
fNoA	−.060**	−.045**	−.088**
bNoA	−.273**		
rNoA		−.172**	
R^2	.447**	.422**	.412**
AoA	.160**	.189**	.223**
R^2	.457**	.438**	.432**
ΔR^2	.010**	.016**	.020**

zRTs: z-score transformed reaction times; fNoA: forward number of associates, bNoA: backward number of associates; rNoA: reciprocal number of associates; AoA: age of acquisition.
** $p < .01$

Of primary importance, the semantic locus hypothesis predicts the first four of these results and is agnostic towards the last. According to the hypothesis, new words (e.g., *apex*) are learned, in part, through associations to previously learned words (e.g., *top*; Steyvers & Tenenbaum, 2005). This premise leads to the prediction that there should be a negative correlation between AoA and bNoA, which was previously established (De Deyne & Storms, 2008). While we knew about the relationship between bNoA and AoA, we did not know how these variables related to reading aloud and lexical decision performance. The results reported here make these relationships clearer. Based on the R^2 values when AoA and fNoA are included in the regression model versus when bNoA is also included, we estimate that about 50% of the previously reported AoA effect in lexical decision RT and about 30% of the AoA effect in reading aloud RT can be attributed to bNoA. However, based on our analyses, AoA still accounts for a significant 1% of additional variance in lexical decision RT and .7% of additional variance in reading aloud RT above and beyond all of the control variables. Unfortunately, we are unable to fully distinguish between bNoA and rNoA as these variables were highly correlated. However, in all cases, bNoA was a stronger predictor than rNoA.[2]

These results are consistent with predictions made by the semantic locus hypothesis (Steyvers & Tenenbaum, 2005), thus advancing our understanding of AoA effects and allowing for a more precise description of the relationship among AoA, bNoA, and performance. In addition, these results have implications for computational models of word recognition. For example, both Multilink and the PDP model of Chang et al. (2019) have incorporated semantic systems into their computational frameworks. Of course, simulations would need to be conducted to determine the extent to which these NoA variables affect the models' performance. However, we think that rNoA may influence word processing performance via interactive activation between the semantic level and the ortho- graphic and phonological levels.

A general approach that accounts for the most basic AoA effects in the literature is the PDP model. A PDP model involves an interconnected network of simple processing units that learn associations between inputs and outputs. Seidenberg and McClelland (1989) introduced what is now becoming the triangle model (for a recent version of this model, see Chang et al., 2019). Triangle models consist of representational units at the orthographic, phonological, and semantic levels. Any one of these levels can serve as an input or output. For example, in a reading aloud task, the orthographic layer serves as the input, and the phonological layer serves as the output. In a visual lexical decision task, orthography serves as the input, and semantics may serve as the output (Cheyette & Plaut, 2017). In between each representation layer is a layer of hidden units. These units mediate associations between the inputs and outputs and allow for arbitrary associations to be formed (Hinton & Shallice, 1991). AoA has a general effect on processing in PDP models, in part because there are larger changes in connection weights early in a model's development than later. This relationship between the amount of weight change and AoA occurs because the amount of weight adjustment is, in part, a function of how much error there is in the output; outputs with more error produce larger weight changes than outputs with less error. Thus, words encountered earlier in training are associated with more error and thus tend to produce larger weight changes than those occurring later. Moreover, the model is exposed to earlier acquired words early in the model's development when the model is considered to be a more optimal processor, thus also contributing to the effect of AoA on word processing (see

Monaghan & Ellis, 2010). Importantly, in PDP models, AoA exhibits the strongest effect when the relationship (i.e., mapping) between inputs (e.g., orthography) and outputs (e.g., phonology) is arbitrary (Chang et al., 2019; Lambon Ralph & Ehsan, 2006, although see Monaghan & Ellis, 2010). When the relation- ship is systematic, later acquired words benefit from the shared structure that they have with earlier acquired words. However, when the relationship is arbitrary, later acquired words do not benefit from shared structure with other words. Consider the process of reading aloud. When the model encounters *baulk* at a later age, there is little cost of being encountered late because of the shared structure between *baulk* and earlier acquired neighbours such as *walk* and *talk*. In contrast, when the model encounters *yacht*, there is less shared structure with other words, and there is a greater processing cost due to its irregular spelling-to-sound pattern. So, the model predicts an interaction between AoA and spelling-to-sound consistency on reading aloud RT (Cortese & Schock, 2013) where inconsistent words are read aloud more slowly when they are later acquired than when they are earlier acquired, and AoA has little effect on words with consistent spelling-to-sound mappings. Interestingly, one of the main predictions of this mapping hypothesis is that there should be a larger effect of AoA on lexical decision RT than reading aloud RT. In lexical decision, if one assumes that the input is orthography and the output is semantics, then the relationship between the two is largely arbitrary. For example, although *baulk* shares structure at the orthographic and phonological levels with *walk, talk*, and so on, it does not at the semantic level. In other words, *alk* maps onto different meanings in these different words. In contrast, in reading aloud, the semi-regular relationship between orthography and phonology in general characterises the inputs and out- puts. In other words, *alk* maps onto the same phonology in all cases. Thus, the PDP approach can explain much of the AoA effects (and how they differ across tasks) and the relative strengths of these effects that are reported in the literature. And, the independent AoA effects reported in the present analyses could be interpreted in a similar manner although the magnitude of the AoA effect is reduced in all cases by the inclusion of bNoA.

So, how would the PDP model explain the bNoA effects reported here? When encountering a word in a reading aloud or lexical decision task, a set of distributed ortho-graphic inputs activates a set of distributed phonological representations and a set of semantic representations via connections mediated by hidden units. Through interactive activation, activated semantic representations activate both phonological and orthographic units, and phonological representations provide input into semantics and orthography. The question here is, in terms of semantic activity, do target words like *top* produce more and/or stronger activation of semantic representations because of their high number of backward associations? It might be safe to assume that the activation of representations associated with the reciprocal associates may occur through recurrent interactive activation among semantic units. In other words, the representations associated with the target are activated and those of the target's associates are then also activated, and this co-activity would serve to reinforce each other. What is not clear is how the representations associated with backward but not forward associates of the target word could also be activated. In the case of reading aloud, the phonological code used for reading aloud typically would be computed quickly and so it would be expected to show smaller semantic effects than lexical decision. In lexical decision, the semantic code would be the best way to differentiate between words and nonwords, and so larger semantic effects would be expected. Alternative computational models of reading aloud and lexical decision (Coltheart et al., 2001; Perry et al., 2010) have not implemented a semantic system. If we assume that the semantic system represents concepts as nodes in a network with forward and backward connections, then the types of interactions between the semantic system and orthographic and phonological systems described for the PDP model would hold for these models as well.

We note that fNoA would be expected to have an effect at the semantic level in terms of increased activity, and this could be used to facilitate lexical decisions. However, one would not expect there to be strong connections between the semantic representations activated by the target's for- ward associates and those semantic, orthographic, and/or phonological representations of the target unless the for- ward associates are also backward associates of the target. Thus, as we observed, fNoA's relationship with lexical decision was reliable, but with reading aloud, the effect of fNoA was much weaker and was only significant when neither bNoA nor rNoA was controlled. We note that Yap et al. (2011) found in their analyses of a smaller set of 389 words that fNoA was not significantly related to either lexical decision or reading aloud performance.

In sum, we differentiated between AoA and bNoA and determined their relationships with RT in reading aloud and lexical decision tasks. These variables can both serve as semantic predictor variables in studies of word processing. We found that both variables affect RTs on both tasks. These effects were more evident in lexical decision—a more semantic-focused task—than reading aloud. In both tasks, AoA effects were reduced when bNoA was con- trolled. Thus, when the semantic locus is operationally defined in terms of the number of backward associates, it provides a viable account of how the AoA of a word seeds semantic representations and influences performance. We also described how contemporary models of word processing may account for these findings, but simulations would need to be done to verify these hypotheses.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Michael J Cortese https://orcid.org/0000-0002-5398-570X

Data accessibility statement

The data from the present experiment are publicly available at the Open Science Framework website: https://osf.io/z8yq6/.

Notes

1. As can be seen in Table 2, orthographic length and Levenshtein

orthographic distance (LOD) were highly correlated. To avoid issues related to collinearity, we included orthographic length and excluded LOD in our analyses. However, none of the patterns regarding AoA or any of the NoA variables change as a function of which variable was included in the model.

2. Researchers may wish to employ structural equation modelling (SEM) to examine related issues (see, for example, Lewis & Vladeanu, 2006). For example, one might structure a model such that bNoA or rNoA, frequency trajectory, AoA estimates, and so on, map onto a latent variable, and this latent variable would then relate to word processing performance. However, we do not think that SEM can differentiate between bNoA and rNoA, as SEM does not solve the problem of collinearity (Lewis & Vladeanu, 2006). Therefore, we did not apply SEM in the present study.

References

Balota, D. A., Cortese, M. J., Sergent Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition for single syllable words. *Journal of Experimental Psychology: General*, *133*, 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Brysbaert, M. (2017). Age of acquisition ratings score better on criterion validity than frequency trajectory or ratings “corrected” for frequency. *Quarterly Journal of Experimental Psychology*, *70*(7), 1129–1139. https://doi.org/10.1080/17470218.2016.1172097

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of familiarity and age of acquisition survive better frequency norms? *Quarterly Journal of Experimental Psychology*, *64*, 545–559. https://doi.org/10.1080/17470218.2010.503374

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, *104*, 215–226. https://doi.org/10.1016/ S0001-6918(00)00021-4

Carroll, J. B., & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior*, *12*, 563–576.

Chang, Y., Monaghan, P., & Welbourne, S. (2019). A computational model of reading across developments: Effects of literacy onset on language processing. *Journal of Memory and Language*, *108*, Article 104025. https://doi.org/10.1016/j. jml.2019.05.003

Cheyette, S. J., & Plaut, D. C. (2017). Modeling the N400 ERP component as transient semantic over-activation within a neural network model of word comprehension. *Cognition*, *162*, 153–166. https://doi.org/10.1016/j.cognition.2016.10.016

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Methods and Research, Instrumentation, & Computers*, *36*, 384–387. https://doi. org/10.3758/BF03195585

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 2340 words. *Quarterly Journal of Experimental Psychology*, *60*, 1072–1082. https://doi.org/10.1080/17470210701315467

Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods*, *40*, 791–794. https://doi.org/10.3758/BRM.40.3.791

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *Quarterly Journal of Experimental Psychology*, *66*, 946–972. https:// doi.org/10.1080/17470218.2012.722660

De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods*, *40*, 213–231. https://doi.org/10.3758/BRM.40.1.213

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M.,

& Rekke, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/ S1366728918000287

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *11*(3), 662–720. https://doi.org/10.1037/0033-295X.111.3.662

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, *98*, 74–95. http://dx.doi.org/10.1037/0033-295X.98.1.74

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, *131*, 684–712. https://doi.org/10.1037/0033-2909.131.5.684

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age of acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*, 978–990. https:// doi.org/10.3758/s13428-012-0210-4

Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition*, *13*, 928–948. https://doi. org/10.1080/13506280544000110

Lewis, M. B., & Vladeanu, M. (2006). What do we know about psycholinguistic effects? *Quarterly Journal of Experimental Psychology*, *59*(6), 977–986. https://doi. org/10.1080/17470210600638076

Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language*, *63*, 506–525. https://doi.org/10.1016/j.jml.2010.08.003

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/

Perry, C., Ziegler, J., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive*

Psychology, *61*, 106–151. https://doi.org/10.1016/j.cogpsych.2010.04.001

Plaut, D., McClelland, J., Seidenberg, M., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115. https://doi.org/10.1037/0033-295X.103.1.56

Schock, J., Cortese, M. J., & Khanna, M. M. (2012). Imageability ratings for 3,000 disyllabic words. *Behavior Research Methods*, *44*, 374–379. https://doi.org/10.3758/s13428-011-0162-0

Schock, J., Cortese, M. J., Khanna, M. M., & Toppi, S. (2012). Age of acquisition ratings for 3,000 disyllabic words. *Behavior Research Methods*, *44*, 971–977. https://doi.org/10.3758/s13428-011-0162-0

Seidenberg, M. S., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568. https://doi.org/10.1037/0033-295X.96.4.523

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. https:// doi.org/10.1207/s15516709cog2901_3

Van Loon-Vervoon, W. A. (1989). *Eigenschappen van basis-woorden*. Swets & Zeitlinger.

Yap, M., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*, 502–529. https://doi.org/10.1016/j.jml.2009.02.001

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review*, *18*, 742–750. https://doi.org/10.3758/s13423-011-0092-y

Zevin, J. D., & Seidenberg, M. S. (2004). Age of acquisition effects in reading aloud: Test of cumulative frequency and frequency trajectory. *Memory & Cognition*, *32*, 31–38. https://doi.org/10.3758/BF03195818