Title: Age-of-acquisition effects: A literature review

Elsherif, M.M.[1] *, Preece, E. [1] and Catling, J.C.[1]

[1] School of Psychology, University of Birmingham

*Address correspondence to:

Mahmoud Medhat Elsherif

52 Pritchatts Road

School of Psychology

University of Birmingham

UK

B15 2TT

Mahmoud.medhat.elsherif@gmail.com

Telephone: NA

Acknowledgements

We thank Magdalena Grose-Hodge and Vikki Anderson for proofreading the manuscript. We thank the three reviewers: Dr Nicolas Dirix, Dr Robert Davies and Professor Mike Cortese and especially, to the editor Dr Tessa Warren for providing critical and constructive feedback that improved our manuscript.

Abstract

Age of acquisition (AoA) refers to the age at which people learn a particular item and the AoA effect refers to the phenomenon that early-acquired items are processed more quickly and accurately than those acquired later. Over several decades, the AoA effect has been investigated using neuroscientific, behavioural, corpus and computational techniques. We review the current evidence for the AoA effect stemming from a range of methodologies and paradigms, and apply these findings to current explanations of how and where the AoA effect occurs. We conclude that the AoA effect can be found both in the connections between levels of representations and within these representations themselves, and that the effect itself occurs through the process of the distinct coding of early and late items, together with the nature of the connections between levels of representation. This approach strongly suggests that the AoA effect results from the construction of perceptual-semantic representations and the mappings between representations.

Keywords: age-of-acquisition; word frequency; word recognition; word production

The Age of Acquisition (AoA) effect or the order of acquisition (OoA) effect[1] was first investigated by Rochford and Williams (1962a, 1962b). Rochford and Williams reported that the number of pictures named by 80% of children aged between 2 and 11 years of age predicted the number of items correctly recognised and produced in aphasic patients. They also found that the lexical items that children found difficult to produce were the most challenging for aphasic patients. This was further demonstrated by Carroll and White (1973), who observed that objects acquired earlier in life, (from hereon referred to as early-acquired) were named more quickly than those which were learnt later in life (from hereon referred to as late-acquired). This effect was also found to be a stronger predictor of naming latencies than word frequency. Over the past six decades, researchers have examined the role of AoA in the processing of words, phrases, pictures, faces and other non-linguistic stimuli (e.g. Arnon et al., 2017; Anderson, 2008; Baddeley & Logie, 1988; Bonin et al., 2008; Carroll & White, 1973; Cortese & Khanna, 2007; Gilhooly & Logie, 1982; Lima et al., 2021; Marful et al., 2018; Morrison & Ellis, 1995; Sereno & O'Donnell, 2009; Smith-Spark & Moore, 2009; Smith-Spark et al., 2012, 2013; Stewart & Ellis, 2008). It is, therefore, well documented that early-acquired items tend to be processed significantly faster and more accurately and are more resilient to forgetting, interference and cortical damage than late-acquired items (e.g. Bonin et al., 2004; Catling & Johnston, 2006c; Cuetos et al., 2010; Juhasz & Rayner, 2003, 2006; Morrison & Ellis, 1995). However, over a decade has passed without a detailed and comprehensive review of the AoA effect, with the four most recent reviews (Brysbaert & Ellis, 2016; Hernandez & Li, 2007; Johnston & Barry, 2006; Juhasz, 2005) discussing different areas of the AoA effect. Thus, in this paper, we aim to provide a comprehensive review of AoA effects on lexical processing. The article provides an in-depth consideration of the underlying mechanisms behind the AoA effect. This is followed by a review of the design of studies and measures used within AoA research, and provides a summary of conclusions from previous reviews. Finally, prior to drawing conclusions and identifying some of the current issues and potential avenues for future

[1] The AoA effect is sometimes called the Order of acquisition effect, as the AoA effect is a measure of the relative order that one acquires an item (Kuperman et al., 2012). This paper will use the term AoA effect, as it is more widely known in the psycholinguistics field (Castles et al., 2018).

research, the paper also addresses the contemporary evidence for the AoA effect in neurotypical participants.

Theoretical Account of AoA effects

Why do early-acquired words have an eminent status in the mental lexicon? The four most recent reviews (Brysbaert & Ellis, 2016; Hernandez & Li, 2007; Johnston & Barry, 2006; Juhasz, 2005) and the current review conclude that AoA effects are likely to have three potential sources: first, due to plasticity mechanisms within the connections between levels of representations (between perceptual/orthography, semantic and/or phonology), second, through the levels of connectivity within the semantic system, , and third, competition between concepts within the semantic system, such that a unique concept must be chosen from its competitors or several word candidates).However, within this review we have taken the second source and subsumed it under a 'multiple loci' account such that the AoA effect is situated in more than one specific linguistic system (i.e. perceptual/orthographic, semantic and phonological). It should be noted that these three sources can be viewed as complementary, as opposed to, contradictory (Brysbaert and Ellis, 2016).

The Multiple Loci Account

The multiple loci account (Catling & Johnston, 2006a, 2006b, 2006c; Moore et al., 2004; Räling et al., 2016) argues that the AoA effect is located in more than one specific linguistic structure (perceptual/orthographic, semantic and phonological). Moore and Valentine (1998) assessed the AoA effects in a familiarity decision task to establish the perceptual processes of the AoA effect. They argued that if an AoA effect is noted during face processing in familiarity decisions, then the AoA effect must have a pre-phonological role, as names are not obligatorily accessed during these tasks. According to Moore et al. (2004), therefore, a single level could not explain the advantage of early acquired words. Instead, they suggest that two separate levels must be involved (phonological and the perceptual level of representations). This is substantiated by findings of studies that have shown AoA effects in tasks that assess verbal output and perceptual input (e.g. Moore et al., 1999, 2004; Moore & Valentine, 1998; Smith-Spark et al., 2012, 2013). These involve motor output tasks, such as picture naming tasks (Bonin et al., 2001; Morrison & Ellis, 1995), which also require access to phonological and possibly orthographic output, and perceptual input tasks, such as visual degradation for pictorial and word stimuli, which involve word and pictorial identification prior to accessing the semantic representation. This has been reliably observed in several studies indicating that the AoA effect is located within more than one linguistic system (e.g. perceptual and phonological; Catling et al., 2008; Dent et al., 2007; Moore & Valentine, 1998).

Catling and Johnston (2006a, 2009) took the multiple loci account a step further and proposed the accumulation hypothesis, according to which the magnitude of the AoA effect accumulates with each additional level of processing necessitated for a specific task. For instance, they showed a smaller AoA effect for word-picture verification than picture naming. They hypothesised that this increase is due to the latter only entailing structural and semantic processing, while the former includes these levels with the addition of phonological processing. Similar to Moore et al. (2004), Catling and Johnston (2009) asserted there were three loci of the AoA effects: perceptual/structural levels, and phonological processing with or without semantics between perceptual and phonology depending on the task.

Representation Theory

Brysbaert and Ghyselinck (2006) argued that the AoA effect is likely to have two components: one of which is linked to word frequency. Brysbaert and Ghyselinck observed that in tasks such as word naming, Lexical Decision Task (LDT) and semantic categorisation, the magnitude of frequency effect and the AoA effect are related to a similar extent: either both effects are small or large. Brysbaert and Ghyselinck concluded that the AoA effect in these tasks is frequency-dependent and occurs at the perceptual or phonological word form level (i.e. phonological/orthographic input lexicon). According to Brysbaert and Ghyselinck, frequency-independent AoA effects result from competition in the semantic/conceptual system (i.e. a unique concept must be chosen from competitors or several word candidates, leading to larger AoA effects than frequency effects, even if rated frequency is larger than rated AoA in semi-factorial and factorial designs). In addition, the representation theory argues that early-acquired words have stronger lateral inhibition that reduces activation of competitors compared to late-acquired words, thus late-acquired words have to expend further effort to overcome their earlier-acquired competitors (Belke et al., 2005). This occurs because there is a correlation between when a new semantic node enters the network and the number of connections that the node will have in the end. In the semantic system, compared to late-acquired words, early-acquired words have more connections to other words and semantic concepts, leading to richer semantic representations (Steyvers & Tenenbaum, 2005). In addition, Steyvers and Tenebaum (2005) argue that the search in lexical retrieval is biased towards accessing early-acquired words first, as they have more highly connected nodes. Steyvers and Tenenbaum demonstrated that AoA effects could be interpreted in light of the richness of semantic connections that early-acquired words possess in the mental lexicon[2]. This, in turn, makes early-acquired words better protected from cognitive impairment (see review by Brysbaert & Ellis, 2016).

The Mapping Theory

According to the mapping theory (Ellis & Lambon-Ralph, 2000; Hirsh & Funnel, 1995; P.Monaghan & Ellis, 2010; Smith et al., 2001), the AoA effect is a property emerging from a learning system and is located primarily in the connections between levels of representation (i.e. between perceptual/orthography, semantic and/or phonology), as opposed to one isolated system such as the semantic system. This theory has been heavily influenced by Parallel Distributed Processing (PDP) theories of cognition. One advantage of this approach is that many of the predicted empirical outcomes can be tested via simulations. Using a connectionist neural network, Ellis and Lambon Ralph (2000) demonstrated that prior to early-acquired items entering the mental lexicon, the neural network has a high level of plasticity. This plasticity benefits early-acquired items, leading to rich and stable representations that are better consolidated in the mental lexicon. Subsequently, the connections between input and output representations are modified by early-acquired items, causing the network to lose plasticity. In other words, early-acquired items have a large effect on the final structure of the network. Consequently, late-acquired items become more difficult to consolidate, producing a recognition and production advantage for early-acquired items over late-acquired items. The AoA effect occurs even if the frequency of later-acquired words is greater than that of early-acquired words, making it difficult for the network to overcome this loss in plasticity. However, late-acquired words may have consistent grapheme-phoneme correspondence and thus be able to borrow the 'knowledge' about words with similar orthography-to-phonology mappings, including early-acquired words, therefore, avoiding

[2] This is despite the fact that the network created by Steyers and Tenebaum was not originally created to account for the AoA effect.

penalty for late entry (Monaghan & Ellis, 2002; Lambon Ralph & Ehsan, 2006; Zevin & Seidenberg, 2002).

J. Monaghan and Ellis (2002b) clarified the predictions of the mapping hypothesis further, by training a neural network with 100 early patterns and 100 late patterns. This model had 80 regular input-to-output mappings and 20 arbitrary input-to-output mappings. In the network, some patterns were trained with high-frequency, whereas others were trained with low-frequency. J. Monaghan and Ellis found an interaction of AoA and consistency, concluding that the AoA effect is more likely to be observed when input-output mapping is arbitrary (between semantics and phonology as seen in picture naming), than when the mapping is systematic and regular (i.e. between orthography and phonology, as shown in word naming). Using connectionist models, Zevin and Seidenberg (2002) replicated these findings and concluded that when the mapping is arbitrary, a genuine AoA effect is demonstrated but AoA effects are merely cumulative-frequency effects when the mapping is systematic and regular.

Extending Zevin and Seidenberg's (2002) model, P. Monaghan and Ellis (2010) incrementally presented words based on the age of the reader with an increasing number of presentations to simulate reading development. P.Monaghan and Ellis replicated the findings of Zevin and Seidenberg that the AoA effect is more likely shown in arbitrary than quasi-regular mapping, but is likely to be shown in quasi-regular mapping when cumulative-frequency is controlled. They argued that the AoA effect is not a cumulative-frequency effect when the mapping is systematic and regular. The AoA effect is not only, therefore, limited to systems involving arbitrary mapping but also quasi-regular mapping. In addition, P. Monaghan and Ellis found that early-acquired words modified the connections of the neural network, leading to rich and stable representations, thus benefitting early-acquired words. Early-acquired words without one-to-one correspondence between graphemes and phonemes (present in languages with deep ortography, such as English) produce a greater change in terms of plasticity since there is competition to overcome in terms of pronunciation. This leaves little plasticity for late-acquired words, making it more difficult for them to consolidate and hence the penalty for late entry. In other words, the AoA effect should be demonstrated in all tasks but AoA effects should increase when the mapping between input and output is arbitrary, as opposed to systematic and regular. To sum up, the mapping theory argues that the AoA effect results from reduced neuroplasticity during the learning of mappings between representations over time and that early plasticity influences processing.

Cautionary tales of AoA and word frequency

Before discussing the evidence of AoA effects, it is important to consider the quality of the AoA and frequency measures that are used within these studies. The measures used to assess AoA and word frequency are not always precise and reliable estimates of these variables, which affects the reliability of findings. Originally, AoA researchers such as Gilhooly and Logie (1980) asked adults to recollect the age at which words were acquired, using a 7-point scale. Subsequently, the reliability of adults' reports of the age of acquisition for different linguistic items has been questioned. In order to check the reliability of the measure, Morrison et al. (1997) used a large norming study in which they collected AoA ratings from young adults and compared the ratings to the objective measures (e.g. naming pictures) obtained from children aged 2 years and 6 months to 10 years and 11 months. Morrison et al. reported that there was a strong correlation (r = .75) between young adult AoA ratings and measures of objective AoA that were recorded from object naming in children. This is not limited to only British and American culture, and has been replicated and generalised (r = .50 - .96) in several languages: Chinese, French, Icelandic, Italian, Kannada, Spanish and Turkish (Álvarez & Cuetos, 2007; Bangalore et al., 2022; Chalard et al., 2003; Chedid et al., 2019; Göz et al., 2017; Li et al., 2022; Liu et al., 2011; Lotto et al., 2010; Montefinese et al., 2019; Perez & Navalon, 2005; Pind et al., 2000; Pind & Tryggvadóttir, 2002). Recently, these AoA measures have also been found to be correlated (r = .29-.91) between 35+ languages, including Arabic, Chinese, Gaelic, English and Persian for pictures of objects and actions (Łuniewska et al., 2016, 2019), indicating the reliability and generality of adult ratings for AoA across cultures and languages. It is important to note that compared to adult ratings, objective AoA has been found to be a better predictor of naming speed (e.g. Pind & Tryggvadóttir, 2002), indicating that objective AoA measures produce stronger AoA effects than adult ratings. These studies show that in the absence of objective AoA masure, adult ratings are sufficiently reliable.

Most measures of AoA are obtained using Likert scales. However, it has been reported that participants find it easier to write when they acquired a word rather than indicating the age on a scale (Kuparman et al., 2012). Another limitation of these studies is that small sample sizes are often not representative of the population (e.g. participants with similar educational attainment only). An example of a study without these limitations is Kuperman et al. (2012). Kuperman et al. recruited more (demographically) diverse sample in terms of age (15 to 82 years) and educational background (a certificate of primary education to a postgraduate degree) in a more natural environment, using Amazon MTurk, and avoided Likert scales to rate 30000 English words for AoA. Participants were asked to type the number corresponding to the age at which they had learned a given word (e.g. type number 3 if a word was acquired at the age of 3) and N or X if they did not know the word at the time of data collection. The results showed that neither the level of education nor the age of participants contribute to AoA. Furthermore, Kuperman et al. observed that self-reported AoA explained 4% and 10% of the lexical decision latencies and lexical decision accuracy respectively, once word frequency, word length, syllables and similarity to other words were considered. This indicates that the age when words are acquired is not a result of educational or age confounds, and that adult ratings can assess the age at which a word is acquired reliably.

It is well evidenced that adults provide good estimates of when early childhood words are acquired, despite remembering little or nothing about these times (i.e. childhood amnesia; e.g., Bauer et al., 2007; Pathman et al., 2013). Nevertheless, one suggestion to improve the validity of the subjective ratings of AoA of these estimates is to assess rater characteristics including how much experience one has with preschool children (Barrow et al., 2019). For example, Barrow et al. investigated the validity and reliability of subjective ratings of AoA by assessing rater characteristics. They asked preschool teachers and individuals with no experience of working with children to validate the AoA ratings of child speech and adult speech. They found that practitioners who worked with children were more accurate in estimating the exact values of a word's AoA. Furthermore, few AoA studies discuss the exact form of the target questions used to produce the objective AoA ratings. In 35+ languages, Luniewska et al. (2016, 2019) collected participants' ratings and changed the question from "When did you learn this word?" to "When do children learn this word?" and found that participants gave AoA estimates half a year below the score to the latter question. They assessed the correlation between "When did you learn this word?" and "When do children learn this word?" and found that they were highly correlated (r =

.93), indicating the reliability of these ratings across languages, and that participants can approximate the order at which words are acquired.

Another issue raised in relation to AoA effects is their genuineness, as AoA is correlated with other psycholinguistic variables (e.g. frequency, imageability and concreteness; Lewis & Vladeanu, 2006; Smolik & Filip, 2022; Fasquel et al., 2022; Strain et al., 1995; Zevin & Seidenberg, 2004; Taylor et al., 2022; Wang & Chen, 2020). Early-acquired words are assumed to be easy to learn, as they are high in frequency, familiarity and imageability, and the converse is true for late-acquired words; which are perceived as difficult to learn. If this assumption is correct, ratings may not be linked to AoA but rather lexical retrieval in adults, making the ratings prone to problems such as circularity of logic. These confounds of AoA with other measures have raised questions concerning the validity of AoA ratings as an independent variable (Zevin & Seidenberg, 2002). However, it has been demonstrated that, when these possible confounds are controlled for, the AoA effect is still evident, suggesting it cannot be simply reduced to the related variables and has strong construct validity[3] (e.g., Brysbaert, 2017; Brysbaert & Biemiller, 2017; Chang et al., 2019; Cortese & Khanna, 2007; Cortese et al., 2018; R.Davies et al., 2017).

The increase in AoA research has also forced researchers to re-evaluate measures of word frequency (e.g., Brysbaert & Cortese, 2011; Brysbaert & New, 2009). Using a poor measure of word frequency, such as CELEX, has been found to overestimate the influence of the AoA effect, while the SUBTLEX-US reduced its influence (Brysbaert & Cortese, 2011), leading to the conclusion that not all frequency measures are equal (Brysbaert & Cortese, 2011). Norms such as those of Baayen et al. (1995), Kucera and Francis (1967), Francis and Kucera (1982) and Zeno et al. (1995) are drawn from corpora that are limited in size and restricted to dated and complex texts, making it difficult to extrapolate beyond a specific sample of items, and hence difficult to assess the real extent of frequency and AoA impact on lexical tasks. Recent corpora that include other forms of written communication such as subtitles of films and less complex written text outperform measures based solely on spoken and other written corpora in terms of predicting lexical retrieval (Brysbaert et al., 2012; Brysbaert & New, 2009;

[3] Construct validity is the degree that an objective measure is a valid representation of a hypothetical construct (i.e. Parsons et al., 2022).

Ernestus & Cutler, 2015; Herdağdelen & Marelli, 2017; Mandera et al., 2015; New et al., 2007). This does not, however, mean written corpora are superior to spoken ones. The latter are known to be small and restricted in a variety of contexts sampled and topics discussed, thus when working with them, more recent measures need to be used to assess the role of word frequency and AoA in lexical retrieval. In addition, researchers should use country-specific SUBTLEX frequency measures. For instance, in the UK, SUBTLEX-UK predicts better measures of lexical retrieval than SUBTLEX-US, while the converse is demonstrated for the USA (Brysbaert & New, 2009; van Heuven et al., 2014). It is, therefore, important to use the most optimal measures of AoA and word frequency when researching the loci and mechanisms of AoA and frequency.

Although the optimal measures of AoA and word frequency could be argued to be objective measures, there is an inherently subjective component to these measures. Older and young adults have different learning histories, producing cohort effects, thus older adults and young adults encounter words differently. For example, young adults encounter words related to technology such as ‘computer’ early in life and more frequently, while older adults may acquire these later in life and less frequently. Consequently, it is important to obtain AoA and frequency ratings from older and younger adults to better predict the performance of neurotypical adults (Cuetos et al., 2012; Deyne & Storms, 2007) or use several word frequencies to assess the corpus best used for the participant sample tested.

Analyses and Designs of AoA research

Traditionally, investigations into the AoA frequency effects were conducted in small-scale studies with fully-factorial designs. In this type of design, small numbers of items are carefully chosen and compared against each other. For instance, in a 2 x 2 design that compares AoA (early and late) and frequency (high and low), four conditions are being compared. These conditions are usually closely matched on a range of psycholinguistic variables (e.g. imageability, word length) and differ as much as possible on variables of interest.

Factorial design is not optimal as it is impossible to have a good range of values across both variables, which contribute to the mixed findings in the literature. When researchers use narrow AoA and frequency ranges, the results will indicate smaller or no AoA and/or frequency effects and false negatives. For example, in experiment 2 and 4, Turner et al. (1998) divided items into early-acquired words with an average value: 2.39 months, and late-acquired words with an average value: 3.50 months, producing an average AoA effect of 25 ms. On the other end of the spectrum lie studies that have used a broad scale with extreme values, which have overestimated the effects of AoA and frequency, leading to false positives. For example, in experiment 1, Gerhand and Barry (1999a) divided words into early-acquired words with an average value: 2.69 months and late-acquired words with an average value: 4.86 months, producing an average AoA effect of 60 ms. These narrowly small or extreme values may not generalise to a larger population of words. Additionally, in this type of study design, it is difficult to have a range of stimuli that is representative of an individual's wider vocabulary, and a large item pool for frequency and AoA, when controlling for all potential confounding variables such as word length, word frequency, contextual diversity and number of syllables (Cutler, 1981).

It has been recommended that optimally researchers should sample words from across a range of values and possibly multiple psycholinguistic dimensions, collect behavioural responses; and estimate or test associations between variance in responses as well as variation in psycholinguistic properties (hereon defined as a regression design; Lewis, 2006). In this way, the stimuli from the entire range of a variable can be used, as long as the researchers are aware of collinearity between measures and that correlations are not too high (i.e. below 0.6; Brysbaert & Ellis, 2016). In turn, this allows for the more reliable and accurate values for rated AoA on a specific task. In addition, if a frequency or AoA measure is observed not to be a reliable measure of these constructs, then it can be replaced with a superior measure, without violating the design of a factorial experiment. All in all, it is recommended that this approach should be used with a large item sample size incorporating a broad range of values and optimal measures of frequency and AoA to assess the effects.

However, regression design and factorial studies are correlational in nature. Studies need to assign stimuli to any condition at random to draw causal conclusions. This is not possible when measuring word frequency or AoA effects, as one cannot randomly assign words to the AoA and frequency condition. For instance, in a single experiment, if the lexical item '*dragon'* is placed in the early-acquired condition, it will not be possible to place it in the late-acquired condition as well. This makes the factorial design and regression design at best a correlational design when testing the AoA effect

(Lewis & Vladeanu, 2006). In order to perhaps overcome the aforementioned issues, one approach is to use stimuli that are not defined as early-acquired or late-acquired such as checkerboards, pseudo-objects or pseudowords (e.g. Catling et al, 2013; Joseph et al., 2014; Stewart & Ellis, 2008). These items can be presented early in the study for one participant and then, late in another study, thus the items could be randomly assigned to the AoA conditions, allowing for a more 'causal' conclusion to be inferred.

In summary, it is important to consider the range of AoA values that are used to ensure that the AoA effect is not the result of item-specific dimension or a limited range of values. In addition, we should strive to improve statistical power to unravel the AoA effect by increasing the number of items and participants to ensure the effects are not the result of item and participant sampling biases. Furthermore, we should ensure that the corpus used to provide the range of AoA and word frequency is reliable and optimal for the task, since not all corpora are created equally. One approach that can be used is a regression design, where a large range of the AoA values and a large subject and item sample size selection.

Evidence for AoA effects

In this section, the evidence for the AoA effect according to the task is outlined. The review is organised in this way as the three predominant theories of AoA make different predictions regarding the presence and/or magnitude of the AoA effects observed in certain tasks (see Tables 1-10 for details of magnitude of AoA and frequency effects for common psycholinguistic tasks). Whilst reviewing the earlier AoA research (including some of the seminal works), we focused primarily on contemporary studies. We calculated the magnitude of the AoA effect between tasks as the following: we took data from all the published studies that assessed for the AoA effect and frequency effect from 1973 – 2020 and included the task as an independent factor, and the AoA effect as an outcome measure. These predictions are discussed in more detail when the theories are evaluated towards the end of the article.

The multiple loci account predicts that the AoA effect would be present in all tasks, but the magnitude of the AoA effect increases with each additional level of processing being accessed. Put simply, the AoA effect would be smallest in tasks such as visual duration threshold and largest in picture naming tasks (e.g. Catling & Johnston, 2009). The representation theory predicts that AoA should be observed in any task that requires access to semantics and that the AoA effect should be larger in tasks that rely more on access to semantics. In addition, the AoA effect should be observed in any task where a frequency effect is observed. However, this AoA effect should be similar in terms of magnitude to that of the frequency effect in all tasks that do not require selection of a unique lemma. Finally, the mapping theory also predicts AoA effects in all tasks, but these effects should be larger when the mapping from input to output is arbitrary[4].

Picture Naming

One of the most common methods of assessing AoA effects is picture naming. This section summarises the findings from published picture naming studies. Results from several of the more recent picture naming experiments using a regression design and factorial design are presented in Tables 1 and 2, respectively.

Each of the three AoA theories predicts an AoA effect in picture naming. If AoA effects are not observed, this would challenge all the theoretical explanations. Our review found that the AoA effect was observed in all of the studies using picture naming. The findings from picture naming latencies are important for representation theory, especially as it predicts that AoA effects should be larger than word frequency effects in picture naming tasks. In addition, according to all three accounts, the magnitude of the AoA effect in picture naming should be larger than for the lexical decision (LD). The multiple loci account argues that as picture naming includes additional representations being accessed, such as phonology, together with semantics, that there should be greater AoA effects for picture naming than word naming and LDT. In addition, the representation account would argue that more semantic processing is involved in picture naming than the LDT, and thus the AoA effect should be larger. The mapping theory suggests that as picture naming has less regular and systematic mapping between representations (phonology and semantics), the AoA effect will be larger. In contrast, the mapping between representations in LDT is more regular and systematic than in picture naming, resulting in a smaller AoA effect.

[4] The AoA effect has been well documented in recognition memory, however this is beyond the scope of the manuscript (see overview by Macmillan et al., 2021).

Table 1.

Findings using a regression design in spoken picture naming.

Study	Language	Group	N_s	N_i	No of trials	AoA	Frequency	Interaction	Significant predictors	Non-significant predictors
Carroll and White (1973)	English	YA	50	103	5150	✓	**X**		L	L, SL
Morrison et al. (1992)	English	YA	20	48	960	✓	**X**		L	I, PT
Vitkovitch and Tyrrell (1995)	English	YA	16	40	640	✓	**X**		NA	VC
Snodgrass and Yuditsky(1996) [a]	English	YA	78	250	19500	✓	✓		Fam, NA	IA, L
Barry et al. (1997)	English	YA	26	195	5070	✓[b]	✓	✓EA:**X** LA: ✓	NA, IA[b]	VC, Fam, L, I
Ellis and Morrison (1998)	English	YA	30	235	7050	✓	✓		L, Fam, NA, IA	L, CFam[c], I[c]
Cuetos et al. (1999)	Spanish	YA	64	140	8960	✓	✓		L, Fam, NA, IA	VC
Kremin et al. (2001)	French	YA	56	140	7840	✓	**X**		NA	VC, L
Dell'acqua et al. (2000)	Italian	YA	84	266	22344	✓[d]	**X**		PT, Ca, H	L, Fam, NA
Bonin et al. (2002)	French	YA	36	203	7308	✓	**X**	**X**	NA, IA, I	Fam, VC, L
Laws et al. (2002)[e]	English	YA	20	120	2400	✓	**X**		EO, L	Fam, CO, VC
Pind and Tryggvadottir (2002)	Icelandic	YA	23	175	4025	✓	**X**		NA, Fam	IA, L
Bonin et al. (2003)	French	YA	120	299	35880	✓	✓		NA, IA	I, VC, Fam, P
Cuetos and Alija (2003)[f]	Spanish	YA	50	100	5000	✓	**X**		NA	VC, I^{+}, Fam, SL
Cuetos and Alija (2003)[g]	Spanish	YA	50	100	5000	✓	**X**		NA	VC, I^{+}, Fam, SL^{+}
Morrison et al. (2003)	English	YA	44	110	4840	✓	**X**		NA	Fam, VC, IA,L, IP
Morrison et al. (2003)	English	OA	30	110	3300	✓	✓		NA, VC	Fam, IA, L, IP
Bonin et al. (2004)	French	YA	60	142	8520	✓			NA, IA	VC, I, CF, Fam, P, Dur
Alario et al. (2004)[h]	French	YA	46	329	13800	✓	✓		NA, IA, CFam, P, I	VC, SL[i]
Alario et al. (2004)[j]	French	YA	46	329	13800	✓	✓		NA, IA, CFam, VC, I, SL	P
Schwitter et al. (2004)[K]	French	YA	40	112	4480	✓	**X**		NA, IA	Fam, I, SL^{+}
Nishimoto et al. (2005)	Japanese	YA	120	260	31200	✓[l]	**X**		NA	Fam[m], M
Severens et al. (2005)	Dutch	YA	40	590	23600	✓	**X**		NA, AoA, SL	L, P
Weekes et al. (2007)	Chinese	YA	100	232	23200	✓	**X**		NA, Fam	VC, IA, CA, SL
Kauschke and von Frankenberg (2008) [l]	German	YA	31	36	1116	✓[m]			I	P, NA, VC
Kausche and von Frankenberg (2008)[k]	German	YA	31	36	1116	✓			NA	I, P, VC
Johnston et al. (2010)	English	YA	25	544	13600	✓	✓		PNA, Fam, NA	VC, L
Lotto et al. (2010)	Italian	C	300	223	66900	✓	**X**[n]			PT, Fam
Liu et al. (2011)	Chinese	YA	30	435	13050	✓	X		CFam, CA, NA, IA	I, VC, L
Bakhtiar et al. (2013)	Persian	YA	100	200	20000	✓	✓		NA, IA	IP, I, L, VC, Fam

Valente et al. (2014)	French	YA	30	120	3600	✓	**X**		NA, IA	PLD, PsF, Fam, VC
Khwaileh et al. (2014)	Levantine Arabic	YA	22	235	5170	✓	**?**[o]		I	VC[o], NA[o], P[o], G[o], Pt[o], R[o]
Shao et al. (2014) [k]	Dutch	YA	74	104	7696	✓			I, IA, NA, VC	
den Hollander et al. (2019)	Dutch	YA	20	140	2800	✓			AoA*Age	
den Hollander et al., (2019)	Dutch	OA	20	140	2800	✓			AoA*Age	
Navarrete et al. (2013)	Italian	YA	20	360	7200	✓	**X**		NA, Ma	VC
Ramanujan and Weekes (2020) [p]	Hindi	YA	40	154	6160	✓	✓		NA, IA, Fam	N, VC, LAS, IP
Karimi and Diaz (2020)	English	YA	212	1901	403012	✓	✓		NA, SL, VC, PN*AoA, PN*NA	PP, PN, PN*F, PN*F*NA, PN*AoA*NA, PN*AoA*F
Bangalore et al. (2022)	Kannada	YA	35	185	6475	✓	**X**		IA, Fam, NA	VC
Bangalore et al. (2022)	Kannada	OA	33	185	6105	✓	**X**		IA, Fam, NA	VC
Wolna et al. (2022)[l]	Polish	YA	95	168	15960	✓	✓		NA, CFam, I, Cindex	GoD, IA
Wolna et al. (2022)[k]	Polish	YA	95	146	13870	✓	✓		NA, CFam	GoD, IA, I, Cindex

Note. The language investigated; type of group tested; sample size of participants and items; total number of trials; a significant (✓) at p <.05 or non-significant (X) effect of age of acquisition (AoA) and frequency; interaction between these effects; other variables included in the equation. N_s = Number of subjects; N_i=number of items; YA = young adults; OA = old adults; C = children; EA = early-acquired words; LA = late-acquired words; L = word length; I = imageability; PT = prototypicality; NA = name agreement; VC =visual complexity; Fam = familiarity; IA = image agreement; CFam = concept familiarity; EO = Euclidean overlap; CO = Contour Overlap; P = number of phonemes; SL = Syllable length; IP = initial phoneme; CF = cumulative frequency; Dur = duration; M = Mora; CA = concept agreement; PNA = picture-name agreement; PLD = phonological Levenshtein distance; PsF = positional segment frequency; G = gender; Pt = plural type; R = rationality; Ma = manipulability; N = nameability; LaS = length on alpha-syllables; PN = phonological neighbourhood density; PP = phonotactic probability; GoD = Goodness of depiction; Cindex = Complexity index. [a] = This refers to Experiment 1; [b] = A significant effect was only observed, as one-tailed t-test was used; [c]= Only using the Lorch and Myers (1990) procedure, the effects of I and CFam were significant; [d]= Including concept agreement in the regression model, makes the AoA effect non-significant but once removed, the AoA effect was significant; [e] = The findings of Laws et al. (2002) are from the 0% masking condition; [f] = Phase 1 results only with errors being mentioned to participant; [g] = Phase results only without errors being mentioned to the participant; [h] The findings reported here are from the first session; [i] = Once the number of phonemes are included in the regression equation, syllable length was non-significant but once phoneme length was removed, syllable length was significant; [j] = the findings reported here are from the second session; [k] = The results reported here are based on action pictures/verbs; [l]= The AoA effect was significant only when a liberal criterion of name agreement was used; [m] = The effect was only significant when a specific measure of name agreement (H) was used; [l] The results reported here are based on object pictures/nouns; [m] = The study controlled name agreement, age of spontaneous production and written word frequency; [n] = The frequency effect was significant, once the AoA effect was removed from the regression equation; [o] = The results reported is unclear whether the effect contributed to the regression model, but the effect was not put in the regression equation; [p] = These adults were Hindi-English bilinguals; $^+$ .05 < p < .10.

Table 2.

Findings using a factorial design in spoken picture naming.

Study	Type	Language	Group	Type of factorial design	N_s	N_i	No of trials	AoA	Frequency	Interaction	Variables controlled
Ellis and Morrison (1998)	Im	English	YA	SF	20	50	1000	176^*			VC, NA, F, I, L
Barry et al. (2001)[a]	Im	English	YA	SF	24/24	24/24	576/576	92^*	-23[b]		L, Fam, NA, VC, IA
Bonin et al. (2001)	Im	French	YA	SF	30/30	36/34	1080/1040	147^*	10		L, PGC, NA, IA, VC, BF, CFam, I
Morrison et al. (2002)	Im	English	YA	SF	35	50	1750	212^*			L, I, VC, F, NA
Morrison et al. (2002)	Im	English	OA	SF	32	50	1600	150^*			L, I, VC, F, NA
Morrison et al. (2002)	Im	English	E	SF	29	50	1450	167^*			L, I, VC, F, NA

Meschyan and Hernandez (2002)[c]	Im	English	YA	Fac	30	80	2400	115*	31	**X**	L, I, VC, NA, P, SL
Barry et al. (2006)[d]	Im	English	OA	SF	10	24	240	74*			SpF, WF, Fam, NA, IA, VC, Con
Barry et al. (2006)[d]	Im	English	YA	SF	9	24	216	78*			SpF, WF, Fam, NA, IA, VC, Con
Catling and Johnston (2006a)[e]	Im	English	YA	SF	24	48	1152	140*			WF, SpF, KF, Fam, NA, P, IA, VC
Catling and Johnston (2006a)[f]	Im	English	YA	SF	24	48	1152	119*			WF, SpF, KF, Fam, NA, P, IA, VC
Catling and Johnston (2006a)[g]	Im	English	YA	SF	48	48	2304	100*			WF, SpF, KF, Fam, NA, P, IA, VC
Catling and Johnston (2006a)[h]	Im	English	YA	SF	48	48	2304	113*			WF, SpF, KF, Fam, NA, P, IA, VC
Catling and Johnston (2006b)[i]	Im	English	YA	SF	15	48	720	127*			WF, SpF, KF, Fam, NA, P, IA, VC
Catling and Johnston (2006c)[j]	Im	English	YA	Fac	20	56	1120	123*			F, Fam, NA, IA, VC, L
Chalard and Bonin (2006)	Im	French	YA	SF	27	60	1620	94*			NA, IA, CFam, VC, F, L, P, SL, I[k]
Holmes and Ellis (2006)[l]	Im	English	YA	SF	21	50	1050	122*			NA, VC, CFam, I, F, SL, IP
Holmes and Ellis (2006)[m]	Im	English	YA	Fac	25	84	2100	86*			NA, SL, F, OF, PNA, VC
Holmes and Ellis (2006)[n]	Im	English	YA	Fac	25	84	2100	41[b]			NA, SL, F, OF, PNA, VC
Lambon Ralph and Ehsan (2006)	Im	English	YA	Fac	44	80	3520	226*[o]		**X**[o]	VC, NA, L, P, SL
Catling et al. (2008)[p]	Im	English	YA	SF	20/20	42/60	840/1200	20*	46*		PNA, Fam, VC,NA, NAN, L, F/AoA
Catling et al. (2008)[q]	Im	English	YA	SF	20/20	42/60	840/1200	54*	34*		PNA, Fam, VC,NA, NAN, L, F/AoA
Catling and Johnston (2009)[r]	Im	English	YA	SF	24	48	1152	71*			SpF, WF, Fam, NA, P, IA, VC
Laganaro and Perret (2011)	Im	French	YA	SF	20	92	1840	26*			CF, NA, VC, PN, P,So, P, PF, SyF
Raman (2011)	Im	Turkish	YA	SF	15	60	900	50*		**X**	P, L, SL, Fam, VC, IA
Raman (2011)	Im	Turkish	DA	SF	15	60	900	68*		**X**	P, L, SL, Fam, VC, IA
Laganaro et al. (2012)	Im	French	YA	SF	45	120	5400	59*			NA, IA, Fam, VC, CFam, SpF
Perret et al. (2014)	Im	French	YA	SF	21	120	2520	47*			NA, IA, CFam, VC, I, F, CF, L, P, ON, PN, PSF, OSF, GF, PF, So
Preece (2015)	Im	English	YA	SF	22/22	68/68	1496/1496	144*	1		SpF, WF, KF/AoA, Fam, VC, I, PNA, Con, NHD, L, P, SL
Catling and Elsherif (2020)[s]	Im	English	YA	Fac	48	80	3840	65*	3	**X**	AoA: F, VC, NA, L, Fa, I, IA F: AoA, VC, NA, L, IA
Ellis and Morrison (1998)	D	English	YA	SF	20	50	1000	-6			VC, NA, F, I, L

Barry et al. (2001)	D	English	YA	SF	24/24	48/48	1152/1152	-1	-5		L, Fam, NA, IA, VC
Meschyan and Hernandez (2002)	D	English	YA	Fac	30	80	2400	-16	-12	**X**	L, I, VC, NA, P, SL
Holmes and Ellis (2006)[n]	D	English	YA	Fac	25	84	2100	-8			L, VC, F, NA
Catling and Johnston (2009)[t]	D	English	YA	SF	24	48	1152	-2			WF, SpF, Fam, NA, P, IA, VC
Izura et al. (2011)[u]	D	Spanish	YA	SF	24	28	672	-13			VC, Fam, SAoA, WF, P, L
Preece (2015)	D	English	YA	SF	20/20	68/68	1360	-8	-6		SpF, WF, KF/RAoA, OAoA, Fam, VC, I, PNA, Con, NHD, L, P, SL

Note. The type of task involved; the language investigated; type of group tested; type of factorial design; sample size of participants and items; total number of trials; the strength of the age-of-acquisition (AoA) and frequency effect (in milliseconds) and whether these effects and interactions were significant at $p < .05$; and variables that were experimentally or statistically controlled. Semi-factorial studies include the sample size of both participants and items for both the AoA- and frequency-manipulated lists separated by a slash, while other studies are stated to be semi-factorial design as they only measured the AoA effect by itself with no other psycholinguistic predictor being included in the same list (see section on statistical analyses and research design of AoA); Im = immediate naming; D = delayed naming; YA = young adults; OA = older adults; E = Elderly population; SF = semi-factorial design, Fac = Factorial design; N_s = Number of subjects; N_i=number of items; VC = visual complexity; NA = name agreement; F = frequency; I = imageability; L = word length; Fam = familiarity; IA = image agreement; PGC = phoneme-to-grapheme correspondence; BF = bigram frequency; CFam = concept familiarity; P = phoneme length; SL = syllable length; WF = written frequency; Con = concreteness; SpF = spoken frequency; KF = Kucera-Francis frequency; IP = initial phoneme; OF = object frequency; PNA = picture-name agreement; NAN = number of alternative names; PN = phonological neighbourhood; CF = cumulative frequency; So = Sonority of first phoneme; PF = phoneme frequency; SyF = syllable frequency; ON = orthographic neighbourhood; PSF = phonological-syllable frequency; OSF = orthographic syllable frequency; GF = grapheme frequency; SAoA = The AoA of Spanish words; [a] = Findings from Phase 1 of Barry et al. (2001) study; [b] = Significant by participants, not by items; [c] = Findings from the no delay in Meschyan and Hernandez's (2002) Experiment 1; [d] = Findings from Phase 1 of Barry et al.'s (2001) study; [e] = Reporting data of Catling and Johnston (2006a) from unrelated primes in Experiment 1; [f] = Reporting data of Catling and Johnston (2006a) from unrelated primes in Experiment 2; [g] = Reporting data of Catling and Johnston (2006a) from unrelated pictorial primes in Experiment 3; [h] = Reporting data of Catling and Johnston (2006a) from unrelated pictorial primes in Experiment 4; [i] = Reporting data of Catling and Johnston (2006b) from Experiment 2; [j] = Reporting data of Catling and Johnston (2006c) from Experiment 2; [k] = Imageability was significantly different between conditions and was included as a covariate; [l] = Findings of Holmes and Ellis' (2006) Experiment 1; [m] = Findings of Holmes and Ellis' (2006) Experiment 5, in which they factorially manipulated AoA and typicality; [n] = Findings of Holmes and Ellis' (2006) Experiment 6, in which they factorially manipulated AoA and typicality; [o] = The AoA effect was obtained from the discussion and the authors did not state the frequency effect in their paper; [p] = The findings of Catling et al. (2008) reported was based on the no circle overlay condition; [q] = The findings of Catling et al. (2008) was based on the no contrast condition; [r] = The findings of Catling and Johnston (2009) obtained from Experiment 4; [s] = The findings of Catling and Elsherif (2020) obtained from Experiment 1b; [t] = The findings of Catling and Johnston (2009) obtained from Experiment 5; [u] = This was a paired-associate learning task, in which participants had to learn labels based on Welsh words provided and controlled for Spanish AoA in Experiment

Originally, the AoA effect was assessed primarily in English, but it has now been investigated, to a similar degree, in several Semitic, and other alphabetic and logographic languages. Out of the 68 number of picture naming studies in alphabetic languages included in Tables 1 and 2, 34% have observed AoA effects with non-significant frequency effects. This pattern of findings has been replicated across several different language families (e.g. Chinese: Liu et al., 2011; Weekes et al., 2007; Levatine Arabic: Khwaileh et al., 2014; Japanese: Nishimoto et al., 2005). Out of 68 picture naming studies in alphabetic languages included in Tables 1 and 2, 19% have observed AoA effects and frequency effects independently, however, the AoA effect has been found to be larger than the frequency effect. This pattern of results has been replicated in Hindi (Ramanujan & Weekes, 2020) and Persian (Bakhtiar et al., 2013). Based on these findings it can be concluded that AoA is a robust measure for picture naming latencies and that the AoA effect is larger than the frequency effect in this particular task.

However, the reason for the lack of observed frequency effects is unclear. Perhaps the most viable explanation is that word frequency is not as important as traditionally thought. Perret and Bonin (2019) used a Bayesian meta-analysis to investigate which variables are important to control and contribute to naming latencies. They observed that image agreement, name agreement, imageability, conceptual familiarity and AoA all contributed strongly to naming speed, while word frequency on naming latencies was inconclusive and less relevant than traditionally believed (but see Gertel et al., 2020 who observed frequency effects in younger and older adults in picture naming, even when AoA was included as a covariate). Perret and Bonin concluded that overall frequency effects are likely to be shown but the magnitude of the frequency effects may result from differing levels of vocabulary knowledge such that people with high vocabulary knowledge are more likely to show a smaller frequency effect than those with low vocabulary knowledge. One explanation is that more skilled readers have more robust, precise and well-specified lexical representations than less skilled readers, allowing easier and faster access to the mental lexicon (e.g. Brysbaert et al., 2017; Elsherif et al., 2022a, 2022b; Mainz et al., 2017; Perfetti, 2007). This stands in contrast to the AoA effect, as individual differences have been found not to moderate the AoA effect (Hsiao & Nation, 2018; Brysbaert et al., 2017).

The AoA effect in all studies reported in Table 2 is rather large. The average effect from 28 picture naming studies is 107 ms. This AoA effect is substantially larger than that observed in LDT (M = 45ms) (t (49) = 5.24, p < .001, d = 1.47[0.84, 2.11]) and word naming (both discussed later) hence supporting all three accounts of AoA. In addition, the AoA effect is much larger than the frequency effect, supporting the notion of a frequency-independent AoA effect, as suggested by the representation theory.

Word naming

One of the predominant tasks, besides picture naming, that has been used to assess the loci and processes of the AoA effect is visual word naming. This involves a participant reading a visually presented word orally as quickly and accurately as possible (i.e. immediate naming). A variant of the word naming task is the conditional naming task, in which a participant needs to name the word, as opposed to nonwords, while the speeded naming task involves a deadline to name a word. The final task that has been utilised is the delayed naming procedure, which involves a word being presented and the participant waiting until a cue appears to name the word. It has been argued that these tasks reflect access to different processes. For instance, in immediate naming, an effect of regularity is observed, whilst this is not demonstrated in speeded naming (Strain et al., 1995). An effect of imageability and larger effects of word frequency are demonstrated in conditional but not immediate word naming (Cortese et al., 2018). These studies indicate that when performing to a deadline, participants depend more on phonological processes than spelling-to-sound processes, whereas conditional word naming depends more on accessing semantic processing. Delayed word naming only assesses the processes involved in initiating articulation as the word recognition mechanisms are completed. These different tasks are discussed independently but summarised at the end of this section.

According to the mapping theory, AoA effects should be observed in word naming, but the effect should be smaller than for picture naming. In addition, the AoA effect should be observed in immediate, conditional and combinatorial naming for the mapping theory, however, the AoA effect should be larger for irregular English words and opaque orthographies. In the multiple loci account, the AoA effect will be observed in all three tasks but the effect should be larger for conditional naming than for immediate word naming tasks. In the representation theory, the predictions would be that the AoA effect would not be evident in immediate naming, as semantics is not involved in word naming. Kuperman (2013) states that “word naming has been repeatedly shown to be a shallower task in that it does not implicate word semantics” (p.5) and that AoA effects should only be shown in immediate naming if highly imageable words are used, as semantics is more likely to be involved, and that the AoA effect will be smaller for word naming than for LDT. In addition, the size of frequency and AoA effects should be correlated and roughly equivalent.

Immediate word naming

Immediate word naming studies provide tests for the three theories discussed. For example, they allow researchers to assess whether the AoA is a stronger predictor of immediate word naming latency than word frequency. According to representation theory (Brysbaert & Ghyselinck, 2006), they should be equal predictors. However, out of the 54 word naming studies in alphabetic languages included in Tables 3 and 4, 24% have observed AoA effects with non-significant frequency effects, whereas 24% have observed independent AoA and frequency effects. The former pattern of findings has been replicated across several different language families (e.g. Chinese: Chang & Lee 2020; Liu et al., 2007; Japanese: Yamazaki et al., 1997). The majority of the evidence points to AoA independently contributing to naming latencies (e.g. Barry et al., 2001; Chen et al., 2007; Cortese & Khanna, 2007; Dewhurst & Barry, 2006; Hernandez & Fiebach, 2006; Liu et al., 2008; J. Monaghan & Ellis, 2002a, 2002b, but see R. Davies et al., 2017 who used mixed-effect models and observed an interaction between age and AoA effects such that the AoA effect reduces with increasing age, on average), as seen in Tables 3 and 4, 17% have noted frequency effects without AoA effects. For example, Lambon Ralph and Ehsan (2006), who controlled visual complexity and letter length together with letter, phoneme and syllable length, observed no AoA effects in word naming[5]. However, Catling and Elsherif (2020) used the same stimuli as Lambon Ralph and Ehsan and found there was no pattern of AoA and frequency effect but there was an interaction such that AoA effect was present in low, not high, frequency words, whereas frequency effects were demonstrated in late-acquired, not early-acquired, words. Catling and Elsherif argue that the difference between their study and that of Lambon-Ralph and Ehsan is that Catling and Elsherif did not collapse the findings across tasks, whereas Lambon-Ralph and Ehsan did.

[5] It is unclear if word frequency effects appeared in word naming, as they did not report a two-way interaction between frequency and task.

Both authors conclude that the magnitude of AoA and frequency effects depends on the degree to which each variable is manipulated.

There are several arguments that have been raised regarding the absence of the frequency and AoA effects. Gerhand and Barry (1998) pointed out that there were differences in how words were presented across studies, whereas Lambon-Ralph and Ehsan (2006) argued that the lack of frequency and AoA effects may occur as a result of the strength of the manipulations in the stimuli set. Although these are important concerns, Elsherif and Catling (2020) used the same procedures and stimuli set as that of Lambon-Ralph and Ehsan and observed that if the data was split by tasks, different results emerge whereas if the same analytical steps as Lambon-Ralph and Ehsan were observed, then the same results would be noted. This indicates that there are several factors that explain the presence of a frequency or AoA effect in word naming. Nevertheless, these designs used suboptimal measures of frequency such as CELEX[6], were factorial and had a small number of stimuli and participants. A megastudy by Cortese et al. (2018) included AoA, word frequency and several other psycholinguistic properties such as imageability plus interactions between these measures. Cortese et al. observed independent effects of AoA and word frequency in immediate word naming, indicating that the lack of frequency and AoA effects observed in previous studies may have resulted from aforementioned methodological differences not being controlled.

According to the mapping theory, AoA effects should be larger in naming English words with inconsistent spelling to sound. J.Monaghan and Ellis (2002b) manipulated AoA and spelling-to-sound consistency while controlling word length, orthographic neighbours, word frequency and imageability (in a word naming task). They observed independent effects of AoA and consistency but these effects were subsumed under an interaction such that AoA effects were larger in inconsistent than consistent words. This interaction has been replicated in English (J. Monaghan & Ellis, 2002a; but see Strain et al., 2002 and Ellis & Monaghan, 2002), transparent languages such as Italian (Wilson et al., 2012), logographic languages such as Chinese (Chang & Lee, 2020) and megastudies (Cortese et al., 2018). Cortese et al. (2018) observed that the AoA effect interacted with feed-forward rime consistency only

[6] Catling and Elsherif (2020) used SUBTLEX-UK.

in immediate naming, such that the difference between feed-forward consistent words and feed-forward inconsistent words was larger for late-acquired than early-acquired words. This highlights that orthographic transparency moderates the AoA effect, and the finding that the AoA effect is larger for an arbitrary relationship between letter and sound may be an intrinsic property of language processing.

Table 3.

Findings using a regression design in spoken word naming.

Study	Task	Language	Group	N_s	N_i	No of trials	AoA	Frequency	Interaction	Significant predictors	Non-significant predictors
Brown and Watson (1987)	Im	English	YA	28	416	11648	✓	**X**			Fam[+], L[+], IP[+], I, BF, Con, Am
Brysbaert (1996)	Im	Dutch	C	22	204	4488	✓	✓		L	
Yamazaki et al. (1997)	Im	Japanese	YA	26	147	3820	✓	**X**		WAoA, Fam	ChF[+], AP, L, VC
Morrison and Ellis (2000)	Im	English	YA	27	220	5940	✓	✓		IP, L	NHD, Fam, I, IP
Colombo and Burani (2002) [a]	Im	Italian	YA	20	160	3200	**X**	✓		L	RF, CA, Con
Colombo and Burani (2002) [b]	Im	Italian	YA	20	160	3200	✓	✓		L, RF, CA	Con
Morrison et al. (2003)	Im	English	YA	30	267	8010	✓	**X**		IP	Fam, I, L[+]
Morrison et al. (2003)	Im	English	OA	30	267	8010	✓	**X**		IP, Fam	I, L, IP
Cuetos and Barbon (2006)	Im	Spanish	YA	53	240	12720	✓	**X**		L, NHD, IP	CF, Fam, I, FT, SL, SA
Liu et al. (2007)	Im	Chinese	YA	480	2423	1163040	✓	**X**		IP, Reg, Fam, Con, NS, NC, PF, NWF, CF*AoA, CF	IP, HD, AoL, NM
Liu et al. (2008)[c]	Im	Chinese	YA	39	2350[d]	91650	✓	✓		IP, Reg, Con, NS, WAoA, I, AoA*Reg, AoA*Con	IP
De Luca et al. (2008)	Im	Italian	YA	34	120	4080	**X**	✓		L	BF, I, ON, SL
De Luca et al. (2008)	Im	Italian	DA	17	120	2040	**X**	✓		L	BF, I, ON, SL
Izura and Playfoot (2012)	Im	English	YA	20	146[e]	2920	✓	✓		L, ON, I, BF, TF	SL, P
R.Davies et al. (2014) [f]	Im	Spanish	YA	25	2764	69100	**X**	✓		L, NHD, IP	ChiF, I
R. Davies et al. (2017)	Im	English	C-OA [g]	179[h]	160	28960	**X**[i]	✓[j]		IP[k], ON[l], BF[m], Reg[n], I[o]	IP,L, ON[o], BF[p], Reg[+q], I[o]
Cortese et al. (2018)	Im	English	YA	25	2500	62500	✓	✓	**X**	L, NHD, FFRC, FBOC, FBRC, F*I, F*NHD	FFOC, I, F*FFRC F*FFOC, F*L, AoA*FFOC, AoA*FFRC, AoA*I
Cortese et al. (2018)	CN	English	YA	25	2500	62500	✓	✓	✓ HF: X LF: ✓	FFOC, FFRC, FBOC, FBRC, I, F*I, F*L, AoA*FFRC	L, NHD, F*FFRC, F*FFOC, F*NHD, AoA*FFOC, AoA*I
Elsherif et al. (2020)	Im	English	YA	48	236[r]	11328	✓	**X**		IP	IP, LMD[+], L, MF1, MF2, Fa, I, ST, MFa1, MAoA1, MFa2, MI1, MI2, MAoA2

Elsherif et al. (2020)	Com	English	YA	48	236[r]	11328	✓	✓		IP, Fa, LMD, MAoA1, MI1	IP, L, MF1, MF2, MFam1, MFam2, MAoA2, MI2

Note. The type of word naming experiment; the language tested; type of group tested; sample size of participants and items; total number of trials; a significant (✓) at *p* <.05 or non-significant (X) effect of age of acquisition (AoA) and frequency; interaction between these effects; other variables included in the equation that were significant or non-significant; Im = immediate naming; CN = conditional naming; Com = combinatorial naming; DA = dyslexic adults; YA = young adults; OA = old adults; C = children; N_s = Number of subjects; N_i=number of items; HF = high-frequency; LF = low-frequency; Fam = familiarity; L = word length; I = imageability; BF = bigram frequency; Con = concreteness; Am = ambiguity; WAoA = written AoA; ChF = character frequency; AP = number of alternative pronunciations; NHD = neighbourhood density; RF = root frequency; CA = context availability; CF = cumulative frequency; FT = frequency trajectory; SL = syllable length; SA = stress assignment; Reg = regularity; NS = number of strokes; NC = number of components; PF = phonological frequency; NWF = number of word formations; HD = homophone density; AoL = age of learning; NM = number of meanings; TF = trigram frequency; ChiF = children frequency; ON = orthographic neighbourhood; FFOC = feedforward onset consistency; FFRC = feedforward rime consistency; F = frequency; FBOC = feedback onset consistency; LMD = Lexeme meaning dominance; MF1 = first morpheme frequency; MF2 = second morpheme frequency; ST = semantic transparency; MFa1 = first morpheme familiarity; MAoA1 = first morpheme AoA; MFa2 = second morpheme familiarity; MI1 = first morpheme imageability; MI2 = second morpheme imageability; MAoA2 = second morpheme AoA. [a] = The results reported are from nouns; [b] =The results reported are from verbs; [c] = The authors used two methods: a factorial approach and a regression approach, the results reported are from the regression approach; [d] = In the factorial design, 120, as opposed to 2350, words were used; [e] = All items were English acronyms; [f] = Results reported are from Experiment 2 in the Spanish speakers, these results were compared with Barca et al.'s (2002) finding; [g] = The authors used a continuous variable of age and split it in the results section but does not provide participant number for each age group; [h] = The authors do not provide the number of participants following data exclusion for each task, thus the numbers obtained were from the degrees of freedom in Table 4; [i] = A significant AoA was obtained only in children, not young or older adults; [j] = Measured with the SUBTLEX-UK Contextual diversity (Van Heuven et al., 2014); [k] = The strength of the initial phoneme measures decreased with age; [l] = Effect shown in children and older adults; [m]= Effect shown in older adults; [n] = Effect shown in young adults; [o] = Effect not shown in young adults; [p] = Effect not shown in children or young adults; [q] = Effect not shown in children or older adults; [r]= All items were bimorphemic English compound words; [+] .05 < *p* < .10.

Table 4.

Findings using a factorial design in spoken word naming.

Study	Type	Language	Group	Type of factorial design	N_s	N_i	No of trials	AoA	Frequency	Interaction	Variables controlled
Roodenrys et al. (1994)[a]	Im	English	YA	SF	15	16/16	240/240	31*	37[b]		L, Con
Roodenrys et al. (1994)[c]	Im	English	YA	SF	28	28/28	784/784	21*	8		L, Con
Morrison and Ellis (1995)	Im	English	YA	SF	21	48/48	1008/1008	32*	1		L, I
Gerhand and Barry (1998)[d]	Im	English	YA	Fac	30	64	1920	14[b]	22*	**X**	L, I, Con
Gerhand and Barry (1998)[e]	Im	English	YA	SF	30	48/48	1440/1440	32*	23*		L, I
Brysbaert et al. (2000)	Im	Dutch	YA	SF	20	48/48	960/960	11[b]	12[b]		L, I
Barry et al. (2001)[f]	Im	English	YA	SF	24/24	24/24	576/576	32*	9		L
Monaghan and Ellis (2002a)	Im	English	YA	Fac	50	72	3600	13*			F, I, L, ON
Monaghan and Ellis (2002b)	Im	English	YA	SF	20/30	80	1600/2400	17*	14*		F/AoA, I, L, ON
Morrison et al. (2002)	Im	English	YA	SF	28	48/48	1344/1344	57*	14*		L, I
Morrison et al. (2002)	Im	English	OA	SF	32	48/48	1536/1536	29*	-4		L, I
Ghyselinck et al. (2004b)	Im	Dutch	YA	Fac	21	96	2016	17*	9	**X**	L, NHD
Barry et al. (2006)[g]	Im	English	YA	SF	20	24	480	27*			L, F
Havelka and Tomita (2006)[h]	Im	Japanese	YA	SF	20	40	800	102*			F, I
Havelka and Tomita (2006)[i]	Im	Japanese	YA	SF	20	40	800	27[b]			F, I
Hernandez and Fiebach (2006)	Im	English	YA	SF	16	96	1536	13*			L, F, I
Raman (2006)	Im	Turkish	YA	SF	28	50	1400	35*			L, F, I, IP
Dewhurst and Barry (2006)	Im	English	YA	SF	60	64	3840	23*	33*		Con, I, L, IP
Burani et al. (2007)[j]	Im	Italian	YA	SF	30	47/48	1410/1440	4	16*		F/AoA, I, BF, ON, L, SL
Burani et al. (2007)[k]	Im	Italian	YA	Fac	24	64	1280	3	19*	**X**	I[l], BF, ON, L, SL, LS, IP, GPC
B.Chen et al. (2007)	Im	Chinese	YA	Fac[m]	26	56	1456	20*			CF, ChF, PrF, Con, NS
Liu et al. (2008)[n]	Im	Chinese	YA	Fac[o]	39	120	4680	23*			IP, NS, F, I, RAoA
Raman (2011)	Im	Turkish	YA	SF	15	60	900	24*		**X**	P, L, SL, Fam, VC, IA
Raman (2011)	Im	Turkish	DA	SF	15	60	900	43*		**X**	P, L, SL, Fam, VC, IA
Wilson et al. (2012)[p]	Im	Italian	YA	SF	40	78	3120	6+			F, I, Con, IP, L, OC, ON, BF
Wilson et al. (2012)[q]	Im	Italian	YA	Fac[r]	32	88	2816	16*			F, I, Con, IP, L, OC, ON, BF
Wilson et al. (2013)	Im	Spanish	YA	Fac	27	120	3240	7[b]	14*	**X**	L, SL, IP, BF, I, Fam, ON
Wilson et al. (2013)[s]	Im	Spanish	YA	Fac	33	80	2640	14*	18*	**X**	L, SL, IP, NHD, BF, I, Fam[t]
Preece (2015)	Im	English	YA	SF	22	68/68	1496/1496	53*	1		SpF, WF, KF/RAoA, OAoA, Fam, VC, I, PNA, Con, NHD, L, P, SL
Raman (2018)[u]	Im	Turkish	YA	SF	33	50	1650	31*			F, I, IP, L, SL
Raman (2018)[v]	Im	Turkish	YA	SF	34	50	1700	18*			F, I, IP, L, SL
Raman (2018)[x]	Im	Turkish	YA	SF	36	50	1800	15*			F, I, IP, L, SL

Raman (2018)[y]	Im	Turkish	YA	SF	36	50	1800	5			F, I, IP, L, SL
Raman (2018)[z]	Im	Turkish	YA	SF	36	50	1800	5			F, I, IP, L, SL
Raman (2018)[a1]	Im	Turkish	YA	SF	30	50	1500	36*			F, I, IP, L, SL
Raman (2018)[a2]	Im	Turkish	YA	SF	30	50	1500	26*			F, I, IP, L, SL
Raman (2018)[a3]	Im	Turkish	YA	SF	30	50	1500	21+			F, I, IP, L, SL
Raman (2018)[a4]	Im	Turkish	YA	SF	30	50	1500	9			F, I, IP, L, SL
Raman (2018)[a5]	Im	Turkish	YA	SF	30	50	1500	5			F, I, IP, L, SL
Catling and Elsherif (2020)	Im	English	YA	Fac	48	80	3840	5	6	✓ HF: -4 LF: 14* EA: -6 LA: 12*	AoA: F, VC, NA, L, Fa, I, IA F: AoA, VC, NA, L, IA
Morrison and Ellis (1995)	D	English	YA	SF	16	48/48	768	3	-9		L, I
Gerhand and Barry (1998)	D	English	YA	Fac	32	64	2048	-11	-2	**X**	L, I, Con
Brysbaert et al. (2000)	D	English	YA	SF	20	48/48	960/960	7	3		L, I
Ghyselinck et al. (2004b)	D	Dutch	YA	Fac	17	96	1632	-5	14*	**X**	L, NHD
Gerhand and Barry (1999a)	Sp	English	YA	Fac	30	64	27*	26*		✓[b] HF: 20 LF: 35 EA: 18 LA: 33	L, I, Con
Ghyselinck et al. (2004b)	Sp	Dutch	YA	Fac	23	96	2208	14*	9	**X**	L, NHD
Wilson et al. (2014)	Sp	Spanish	YA	Fac	35	120	4200	3	11*	**X**	L, SL, IP, BF, I, Fam, ON

Note. The type of task involved; the language investigated; type of group tested; type of factorial design; sample size of participants and items; total number of trials; the strength of the age-of-acquisition (AoA) and frequency effect (in milliseconds) and whether these effects and interactions were significant at $p < .05$; and variables that were experimentally or statistically controlled. Semi-factorial studies include the sample size of both participants and items for both the AoA- and frequency-manipulated lists separated by a slash, while other studies are stated to be semi-factorial design as they only measured the AoA effect by itself with no other psycholinguistic predictor being included in the same list (see section on statistical analyses and research design of AoA); Im = immediate naming; D = delayed naming; Sp = speeded naming; YA = young adults; OA = older adults; SF = semi-factorial design, Fac = Factorial design; N_s = Number of subjects; N_i=number of items; HF = high frequency; LF = low-frequency; EA = early-acquired; LA = late-acquired; L = word length; I = imageability; Con =concreteness; NHD = neighbourhood density; F = frequency; IP = initial phoneme; BF = bigram frequency; ON = orthographic neighbourhood; SL = syllable length; LS = lexical stress; GPC = grapheme-phoneme correspondence; CF = cumulative frequency; ChF = Character frequency; PrF = phonetic radical frequency; NS = number of strokes per character; RAoA = rated AoA; OC = orthographic complexity; Fam = familiarity; SpF = spoken frequency; WF = written frequency; KF = Kucera-Francis Frequency; OAoA = Objective AoA; VC = visual complexity; PNA = picture-to-name agreement; P = number of phonemes; NA = name agreement; [a] = Data reported from Roodenrys et al.'s (1994) Experiment 2; [b] = Significant by participants, not by items; [c] = Data reported from Roodenrys et al.'s (1994) Experiment 3; [d] = Data reported from Gerhand and Barry's (1998) Experiment 1; [e] = Data reported from Gerhand and Barry's (1998) Experiment 3A; [f]= Findings from Phase 1 of Barry et al. (2001) study; [g] = Findings from Phase 1 of Barry et al.'s (2006) study; [h] =script type (i.e. Kanji and Kana) was a between-participant variable, thus presented as separate groups, data reported from Kanji in Havelka and Tomita's (2006) study; [i] = data reported from Kana in Havelka and Tomita's (2006) study; [j] = The data reported is from Burani et al.'s (2007) Experiment 3; [k]= the data reported is from Burani et al,'s (2007) Experiment 5;[l] = The effect of imageability was included as a covariate because it was significantly different between item conditions but was found not to contribute to naming latencies; [m] = The factorial design included predictability and AoA as a 2 x 2 design such that unpredictable words produced a larger AoA effect than predictable words; [n] = the authors used two methods a factorial and regression approach, the results reported are from the factorial approach; [o] = In the factorial design, 120, as opposed to 2350, words were used; [p]= The authors used bisyllabic words; [q] = The authors used trisyllabic words; [r] = The factorial design included regularity, AoA and block type; [s] = The data is reported from Experiment 4, in which they tested whether the AoA effect would be presented in highly imageable words; [t] = The effect was significantly different between item groups but significantly contributed to naming latencies by item, not subject; [u] = The data is reported from Experiment 1; [v] = The data is reported from Experiment 2 with high-frequency words being used as filler items; [x] = The data is reported from Experiment 3 with moderate-frequency words being used as filler items ; [y]= The data is reported from Experiment 4 with low-frequency words being used as filler items; [z] = The data is reported from Experiment 5 with nonwords being used as filler items; [a1]= The data is reported from Experiment 6 as a replication of Experiment 1; [a2] = The data is reported from Experiment 7 with highly imageable words as filler items; [a3] = The data is reported from Experiment 8 with medium imageable words as filler items; [a4] = The data is reported from Experiment 9 with low imageable words as filler items; [a10]= The data is reported from Experiment 10 as a replication of Experiment 5 with nonwords used as filler items; [+] $.05 < p < .10$.

These findings were limited to monomorphemic and monosyllabic words, thus may not generalise to more complex words. According to the mapping theory, the AoA effects should be larger for complex words such as disyllabic words and compound words. Cortese and Schock (2013) argued that longer and more complex words have less regular letter-sound mapping, thus semantics is more likely to contribute to the naming of such words. In a regression design, Cortese and Schock (2013) assessed the influence of AoA on disyllabic words, while including imageability, frequency and consistency. Cortese and Schock extracted reaction times from the ELP and reported a word frequency effect and an AoA effect, which were subsumed under an interaction such that the AoA effect was smaller in high-frequency than low-frequency words. Juhasz et al. (2015) examined how several predictors such as AoA, imageability and familiarity contribute to naming compound words observing that the AoA, imageability and familiarity of the compound word, not the individual lexeme, contributed to word naming (but see Elsherif et al., 2020, who showed only independent effects of AoA). This indicates conclusively that AoA also contributes to naming complex words and that the AoA effect in naming words is primarily the result of mapping between representations.

To assess the generality of the AoA effect in word naming, it is important to investigate this in other languages. In addition, the mapping theory predicts that the AoA effect should be moderated by orthographic transparency such that transparent orthographies (e.g. Spanish; i.e. letter-sound mapping is more regular and systematic) should produce smaller AoA effects than opaque orthographies (e.g. English; i.e. letter-sound mapping is arbitrary). In contrast to opaque orthographies such as English, where a large AoA effect (i.e. above 20ms) is usually observed, in transparent orthographies such as Italian or Dutch, an AoA effect is present in naming words but the effect is tiny (i.e. around 10ms; Dutch: Brysbaert et al. 2000; Ghyselinck et al. 2004b; Italian: Colombo & Burani, 2002) or the AoA effect is absent altogether (Burani et al., 2007). However, contradictory evidence has shown that two transparent orthographies (i.e. Turkish and Spanish) show AoA effects, as large as those in opaque languages (above 20ms; Cuetos & Barbón, 2006; Raman, 2006). It is unclear as to why the AoA effects were present in Spanish but not Italian. These findings could potentially be explained by semantic involvement occurring in transparent languages, making word reading more opaque. Moreover, the items used by Cuetos and Barbon (2006) and Raman (2006) were highly imageable, whereas Burani et al. (2007) used a varied range of word stimuli, though the average imageability scores were lower than that for words used in Cuetos and Barbon and Raman. High-imageable items have a richer semantic representation and are placed at the centre of the semantic network (Steyvers & Tenenbaum, 2005), encouraging a semantic-mediated process of word recognition. As a result, an AoA effect is more likely to be produced. Using high-imageable items, Raman (2018) and Wilson et al. (2013) found both an AoA effect and a frequency effect. However, when they used low-imageable items, the AoA effect disappeared. R.Davies et al. (2014) compared the results of Cuetos and Barbon, which were superseded in their re-analysis of the original data conducted by R.Davies et al. (2013), using mixed-effect models and Principal Component analysis to deal with collinearity of predictors, and Burani et al., and found that imageability was significantly different between both studies. When imageability was comparable between both languages, the AoA effect shown in Spanish disappeared. This highlights that the previous findings of the large AoA effect in transparent orthographies may well result from the increased level of semantic involvement, making the relationship between orthography and phonology more arbitrary, as semantics is involved as an additional process. These findings cannot be explained by the separate theories alone, but by an integrated account of the AoA effect (see integrated view of the AoA effect later).

Although a large AoA effect has been found in English, results from Dutch, Italian, Spanish and Turkish seem to contradict the mapping theory. Most of these studies were conducted in alphabetic languages. Further evidence in favour of the mapping theory comes from logographic languages such as Chinese and Japanese, which have a more arbitrary print-to-sound relationship, and a more systematic and regular relationship between print-to-meaning. In Japanese, Yamazaki et al. (1997) included character frequency, alternative pronunciations, length, visual complexity, word frequency, spoken AoA, written AoA and familiarity of Kanji nouns in their study. Yamazaki et al. observed that spoken and written AoA, together with familiarity of Kanji nouns, contributed to naming latencies. Interestingly, they observed that written AoA explained more variance in naming latencies of Japanese characters than spoken AoA (but see Morrison, 2003; Shibahara & Kondo, 2002; Yamada et al., 1998 who argue that these findings may be explained by multi-collinearity, using objective, as opposed to rated, AoA and ease of articulation, respectively). Havelka and Tomita (2006) manipulated AoA and script-type as a between-subject variables such that half of the native Japanese speakers were presented with words in Kanji script, a logographic system that has an arbitrary mapping between letter and sound. The Katakana script, a syllabic system with a regular letter-sound mapping, was provided to the other half of the participants. Havelka and Tomita observed that the AoA effect was larger in Kanji (102ms) than Katakana (27ms). The evidence from logographic languages indicates that the AoA effect results from the mapping regularity between print and sound or meaning.

Speeded naming

It has been argued that there should be a larger AoA effect when naming a word under time constraints, as the articulatory lexicon has to be accessed more quickly, while frequency effects should not change, as this would not depend on articulatory processes (Gerhand & Barry, 1999b). Evidence in support of this hypothesis is equivocal (Gerhand & Barry, 1999b; Ghyselinck et al., 2004a). One explanation is that this interaction could result from the degree of manipulation concerning word frequency and AoA (Ellis & Lambon-Ralph, 2000; Lambon-Ralph & Ehsan, 2006). Wilson et al. (2013) used speeded naming to investigate whether AoA effects occur in Spanish and examined whether the AoA effect results from the articulatory lexicon or orthographic-to-semantic mapping. To investigate the latter, a wide range of imageability values were used. Wilson et al. observed a main effect of word frequency but the AoA effect was evident only in highly imageable words, showing that the AoA effect resulted from orthographic-to-semantic mapping in transparent orthographies. The mixed findings may result from the fact that individuals' average naming speeds differ (see also Davies et al., 2017 when discussing individual differences in average word naming speed). As a result, the manipulation of speeded naming ought to consider the baseline speed for the individual to name the word and adjust the speed of naming a word to be an effective manipulation. However, this is not the case, and in order to reduce the ambiguity of this effect, it is important that future research allows for this manipulation to better assess speeded naming.

Delayed naming

Delayed naming is often used to assess whether the AoA effect is present during the initiation of articulation, as the processes underlying word recognition are completed before the onset of the response. Hence any effects resulting from differences in the initial phonemes of the word can be controlled. Several studies have shown no AoA effects on delayed word and picture naming latencies (Barry et al., 2001; Brysbaert, Lange, et al., 2000; Catling & Johnston, 2009; Ellis & Morrison, 1998; Gerhand & Barry, 1998; Ghyselinck et al., 2004b; Holmes et al., 2006; Izura et al., 2011; Meschyan & Hernandez, 2002; Morrison & Ellis, 1995; Preece, 2015). Frequency effects are also rarely shown in delayed naming for the articles included in the current review (but see Balota et al., 1985, 1990; Ghyselinck et al., 2004b, who observed a significant frequency effect). It is important to note that initial phoneme onset is rarely included as a control measure, although it is known to contribute heavily to visual word naming latencies (Treiman et al., 1995). However, when controlling for initial phoneme onset or using delayed naming neither impact the AoA effect (e.g. Elsherif & Catling, 2021; Elsherif et al., 2020; Morrison et al., 1995, 2003).

To summarise, there is evidence that AoA effects are more often observed in picture naming than word naming. Evidence supporting the involvement of semantics is unclear in monomorphemic word naming (e.g. Balota et al., 2004; Elsherif et al., 2020; Cortese et al., 2018) but recent evidence has been converging to show that semantics may not contribute to immediate word naming (Elsherif et al., 2020; Cortese et al., 2018; Kuperman, 2013). According to all three accounts, the AoA effect should be much smaller in word naming than picture naming. Based on the tables, we extracted the means of the AoA effect from picture naming and word naming of each study (excluding delayed naming), and compared them using an independent t-test. We observed that the AoA effect was significant and was about 5 times smaller in word naming (i.e. 22ms) than picture naming (i.e. 104ms) ($t(71) = -9.49$, $p < .001$, $d = 2.26$ [1.66, 2.86]), while the word frequency effect did not significantly differ in the same word naming (i.e. 14ms) and picture naming (i.e. 15ms) tasks ($t(25) = 0.15$, $p = .88$, $d = 0.07$ [-0.79, 0.93], $BF_{01} = 0.34$). This suggests that the AoA effect results from the mapping between representations and is partly lexical-semantic in nature, supporting the findings previously shown by Juhasz et al. (2005) and Brysbaert and Ellis (2016).

Lexical decision

The predictions of the different AoA theories on the LDT are now considered. According to the mapping theory, representation theory and multiple loci account, AoA effects should be observed in

LDT. The multiple loci account argues that as LDT includes additional representations being accessed, such as orthography, together with semantics, there should be larger AoA effect than visual word naming. However, the representation theory argues that more semantic processing is involved in LDT than visual word naming, therefore it is LDT that would be expected to produce a larger AoA effect. In addition, a unique prediction of the representation theory account is that as semantic processing is more likely to be involved, the size of frequency and AoA effects should not be correlated, and the AoA effect should be larger than the frequency effect. Finally, the mapping theory argues that as LDT has less regular and systematic mapping between representations (orthography and semantics) than visual word naming, that the AoA effect will be larger.

Arguably one of the most prevalent techniques for investigating the processes underlying word recognition, the LDT, has repeatedly been shown to be affected by semantic processing (e.g. Balota et al., 2004; Cortese et al., 2018; Schilling et al., 1998). Initially there was some contention about the role of AoA on LDT reaction times (e.g. Barry et al., 2006; Bonin et al., 2001; Brysbaert et al., 2000; Butler & Hains, 1979; Gerhand & Barry, 1999b; Gilhooly & Logie, 1982; McDonald & Shillcock, 2001; Morrison & Ellis, 1995, 2000; Nagy et al., 1989; Schwanenflugel et al., 1988; Whaley, 1978). These findings were based on small-scale studies, with small sample sizes in terms of participants and items, potentially leading to non-robust findings for the AoA effect (see Tables 5 and 6). Using a regression model, a megastudy by Cortese et al. (2018) included AoA, word frequency and other psycholinguistic predictors such as letter length, together with interactions between measures. Cortese et al. observed a main effect of AoA and word frequency in LDT, but these were subsumed in an interaction of word frequency and AoA such that low-frequency words produced larger AoA effects than high-frequency words. This interaction was larger for LDT than immediate word naming, indicating that the lack of frequency and AoA effects in previous studies may have resulted from non-optimal stimuli selection, a limited sample size for items and participants and/or a factorial design.

The AoA effects observed in transparent and opaque languages may result from a strong semantic contribution (Cortese & Khanna, 2007; see also Hsiao and Nation, 2018, who observed that early-acquired words were encountered in varied and diverse contexts and were more semantically related to other known words compared to late-acquired words). Early-acquired words are processed more semantically because they have richer semantic representations and connections, and activate semantic features to a larger extent than late-acquired words (Wilson et al., 2013). Wilson et al. (2013) orthogonally manipulated frequency and AoA in LDT and in word naming. Frequency affected performance in all four experiments within Wilson et al.'s study while AoA effects were observed in the LDT (Experiment 3) and only when naming highly imageable items (Experiment 4). When stimuli contained words from a wider sample of imageability values, AoA did not significantly contribute to the word naming latencies

However, the LDT is also affected by nonword context. Illegal and unpronounceable nonwords are less affected by phonology and orthography (e.g. xycfd), whereas legal and pronounceable word-like nonwords (e.g. haid) are more influenced by additional orthographic and phonological processes (Andrews, 1997). Spataro et al. (2013) used a LDT with illegal nonwords and legal and pronounceable nonwords in the unstudied items. They argued that more word-like nonwords forces participants to process the strings more deeply in order to make a decision. The authors observed independent effects of AoA in word targets. However, they observed that the AoA effects in word targets were larger in legal and pronounceable nonwords than illegal nonwords, as this entails deeper processing (see also Ghyselinck et al., 2004b who replicated these findings in Dutch and Holmes & Ellis, 2006 who observed that AoA effects in object recognition were larger for pseudo-objects that were closer to real objects than non-objects). This indicates that nonword context contributes to the strength of the AoA effect, which is indicative that orthographic and phonological processing both contribute to the AoA effect, along with semantic processing.

Table 5.

Findings using a regression design in lexical decision tasks.

Study	Presentation	Language	Group	N_s	N_i	No of trials	AoA	Frequency	Interaction	Significant predictors	Non-significant predictors
Baumgaertner and Tompkins (1998)	A	English	OA	32	50	1600	✓	✓		Fam[a]	
Nagy et al. (1989)	V	English	YA	95	168	15960	✓	✓		StF, AoA^2, L, DF, DF * PS	PS
Morrison and Ellis (2000)	V	English	YA	24	220	5280	✓	✓		I	L, Fam, IP, NHD[+], PS
Bonin et al. (2001)	V	French	YA	36	237	8532	✓	✓	✓ HF: X LF: ✓	L	Fam, I, BF, NHD, NHF, GPC
Colombo and Burani (2002)[b]	V	Italian	YA	20	160	3200	**X**	✓		CA	RF, L, Con
Colombo and Burani (2002)[c]	V	Italian	YA	20	160	3200	✓	✓		RF	L, CA, Con
Fiebach et al. (2003)	V	German	YA	12	136	1632	✓	✓			
Fiebach et al. (2003)	A	German	YA	14	136	1904	✓	X			
(Boulenger et al., 2007)[b]	V	French	YA	20	153	3060	✓[d]	✓			BF, TF, SL, L,I
Boulenger et al. (2007)[c]	V	French	YA	20	153	3060	**X**	✓			BF, TF, SL, L,I
Deyne and Storms (2007)	V	Dutch	YA	22	108	2376	✓	✓		Fam	
Deyne and Storms (2007)	V	Dutch	OA	20	108	2160	✓	✓			
Menenti and Burani (2007)	V	Italian	YA	54	134	7236	✓	✓		I, L, WC	
Menenti and Burani (2007)	V	Dutch	YA	50	134	6400	✓	✓		L	Fam[+]
González-Nosti et al. (2014)	V	Spanish	YA	36	5530	199080	✓	✓	✓ HF: X LF: ✓	I, ON, AoA*I, F*I, I*NHD, L*NHD, L*AoA, F*L	L
R.Davies et al. (2017)	V	English	C-OA[e]	354[f]	160	56640	✓	✓[g]		IP[h], L, ON[i], I	BF, IP[j], ON[k], R
(Izura and Hernández-Muñoz (2017)	V	Spanish	YA	24	150	3600	✓	✓		L	NHD, Fam, LD, I, NoA, Co

Hsiao and Nation (2018)	V	English	C	114	300	34200	✓	✓	✓ HF: ✓ LF: ✓	L, SD, P[l]	
Cortese et al. (2018)	V	English	YA	25	2500	62500	✓	✓	✓ HF: X LF: ✓	FFOC, FFRC, I, F*FFRC, F*L, F*I	L, NHD, FBOC, FBRC, F*FFOC, F*NHD, AoA *FFOC, AoA*FFRC, AoA * I
Xu et al. (2020)	V	Chinese	YA	1765	19716	34798940	✓[g]	✓	✓ HF: ✓ MF: ✓ LF: ✓	SC, ChF, CF[m], NWF[m], NoP[m]	NoM
Chang and Lee (2020)	V	Chinese	YA	180	3314	596520	✓			ChF, SAR, I	R, Un, SC, Con, PC, SeC

Note. The presentation of the stimuli; language investigated; type of group tested; sample size of participants and items; total number of trials; a significant (✓) at *p* <.05 or non-significant (X) effect of age of acquisition (AoA) and frequency; interaction between these effects; other variables included in the equation. V = visual; A = auditory; N_s = Number of subjects; N_i=number of items; YA = young adults; OA = older adults; C = children; StF = stem frequency; AoA^2 = AoA squared; L = word length; DF = derivational frequency; PS = part of speech; I = imageability; Fam = familiarity; IP = initial phoneme; NHD = neighbourhood density; HF = high-frequency; LF = low-frequency; IV = image variation; BF = bigram frequency; NHF = neighbourhood frequency; GPC = grapheme-phoneme correspondence; RF = root frequency; CA = context availability; Con = Concreteness; WC = word category; ON = orthographic neighbourhood density; R = regularity; LD = Levenshtein distance; NoA = number of associates; Co = connectivity; SD = semantic diversity; P = proficiency; FFOC = feedforward onset consistency; FFRC = feedforward rime consistency; F = frequency; FBOC = feedback onset consistency; MF = medium frequency; SC = Strokes per character; ChF = character frequency; NWF = number of words formed; CF = cumulative frequency; NoM = number of meanings; NoP = number of pronunciations; Un = Unpronounceable; Con = consistency; PC = phonetic combinability; SeC =semantic combinability; SAR = semantic ambiguity rating; [a]= Familiarity disappeared if AoA entered before familiarity in the regression equation; [b] = nouns were used; [c] = verbs were used; [d] = the AoA effect remained when using entire set of items, 2*60 early and late-acquired target nouns and target verbs but trended towards significance in the subset of 152 items; [e] = The authors used a continuous variable of age and split it in the results section but does not provide participant number for each group; [f] = The authors do not provide the number of participants following data exclusion for each task, thus the numbers obtained were from the degrees of freedom in Table 4; [g] = Measured with the SUBTLEX-UK Contextual diversity (Van Heuven et al., 2014) ; [h] = The effect was significant in young adults and older adults; [i] = The effect was significant in children and older adults; [j] = The effect was not significant in children; [k] = the effect was not significant in young adults; [l] = The authors replaced word frequency with document count to assess whether these patterns were artefacts of contextual diversity, the same pattern of results remained; [m] = This was significant in the model based on Tsang et al. (2018), not Kuperman et al. (2012). [+] .05 < *p* < .10.

Table 6.

Findings using a factorial design in lexical decision tasks.

Study	Presentation	Language	Group	Type of factorial design	N_s	N_i	No of trials	AoA	Frequency	Interaction	Variables controlled
Morrison and Ellis (1995)	V	English	YA	SF	16	48/48	768	66^*	54^*		L, I, F/AoA
Turner et al. (1998)	V	English	YA	SF	25/26	66/64	1650/1664	25^*	33^*		L, NHD, I, UP
Turner et al. (1998)	A	English	YA	SF	20/20	66/64	1320/1280	46^*	-8		L, NHD, I, UP
Gerhand and Barry (1999b)[a]	V	English	YA	Fac	20	46	920	59^*	77^*	✓ HF: 10 LF: 109	L,I, Con
Gerhand and Barry (1999b)[b]	V	English	YA	Fac	20	46	920	25^*	33^*	✓ HF: 5 LF: 44	L,I, Con
Gerhand and Barry (1999b)[c]	V	English	YA	Fac	20	46	920	56^*	90^*	✓ HF: 33 LF: 79	L,I, Con
Gerhand and Barry (1999b)[d]	V	English	YA	Fac	20	46	920	39^*	58^*	✓ HF: 22 LF: 33	L,I, Con
Gerhand and Barry (1999b)[e]	V	English	YA	Fac	20	46	920	57^*	66^*	✓ HF: 20 LF: 77	L,I, Con
Brysbaert et al. (2000)	V	Dutch	YA	SF	20	48/48	960/960	52^*	85^*		L, I, F/AoA
Bonin et al. (2001)	V	French	YA	SF	30/30	36/34	1020/1080	56[a]	49^*		F, L, P, SL, ON, NHF, GPC, BF
Ghyselinck et al. (2004b)[a]	V	Dutch	YA	Fac	20	96	1920	75^*	70^*	**X**	L, SL, ON
Ghyselinck et al. (2004b)[b]	V	Dutch	YA	Fac	20	96	1920	12^*	18^*	**X**	L, SL, ON
Ghyselinck et al. (2004b)[c]	V	Dutch	YA	Fac	20	96	1920	117^*	102^*	**X**	L, SL, ON
Barry et al. (2006)	V	English	YA	SF	10	24	240	46^*			L, F, Fam, NA, IA, VC
Barry et al. (2006)	V	English	OA	SF	9	24	216	47^*			L, F, Fam, NA, IA, VC
Burani et al. (2007)[f]	V	Italian	YA	SF	30	47/48	1410/1440	10^*	14^*		F/AoA, I, BF, ON, L, SL

Burani et al. 2007)[g]	V	Italian	YA	Fac	20	64	1280	17*[h]	33*	✓[h] HF: 4 LF: 30	I, BF, ON, L, SL, LS, IP, GPC
Weekes et al. (2008)	V	Chinese	YA	SF	12	44/44	528/528	46*	26*		I, L, SC, F/AoA, SL
Chen et al. (2009)	V	Chinese	YA	SF	32	72	2304	10*			ChF, SC, Con, Fam and H
Sereno and O'Donnell (2009)	V	English	YA	SF	90	100	9000	31*			F, L, GB
Spataro et al. (2013)[a]	V	Italian	YA	SF	22	60	1320	63*[i]			L, WF, Fam, I, Con
Spataro et al. (2013)[b]	V	Italian	YA	SF	22	60	1320	42*[i]			L, WF, Fam, I, Con
Wilson et al. (2013)	V	Spanish	YA	Fac	26	120	3120	31*	57*	✓HF: 15 LF: 47	L, SL, IP, BF, I, Fam, ON

Note. The presentation of the stimuli; The language investigated; type of group tested; type of factorial design; sample size of participants and items; total number of trials; the strength of the age-of-acquisition (AoA) and frequency effect (in milliseconds) and whether these effects and interactions were significant at *p* < .05; and variables that were experimentally or statistically controlled. Semi-factorial studies include the sample size of both participants and items for both the AoA- and frequency-manipulated lists separated by a slash, while other studies are stated to be semi-factorial design as they only measured the AoA effect by itself with no other psycholinguistic predictor being included in the same list (see section on statistical analyses and research design of AoA). V = visual; A = auditory; YA = young adults; SF = semi-factorial design, Fac = Factorial design; N_s = Number of subjects; N_i=number of items; L = letter length; I = imageability, F = frequency; NHD = neighbourhood density; UP = uniqueness point; Con = concreteness; P = phoneme length; SL = syllable length; ON = orthographic neighbourhood density; NHF = neighbourhood frequency; GPC = grapheme-phoneme correspondence; BF = bigram frequency; Fam = familiarity; NA = name agreement; IA = image agreement; VC = visual complexity; IP = initial phoneme; LS = lexical stress; SC = Strokes per character; ChF = character frequency; H = homophone density; GB = gender bias; WF = written frequency. [a] = Nonwords were orthographically legal and pronounceable; [b] = Nonwords were orthographically illegal; [c]= Nonwords were pseudohomophones; [d] this included an articulatory suppression task of reciting Mary had a little lamb; [e] This includes an articulatory suppression with the use of 'the'; [f]= The data reported is from Burani et al.'s (2007) Experiment 4; [g] = The data reported is from Burani et al.'s (2007) Experiment 6; [h] = Significant by participants only; [i] = data is reported from the non-studied list

The findings so far have been limited to monomorphemic words and have not been generalised to morphologically complex words such as disyllabic and compound words. However, Cortese and Schock (2013) used word naming and LDTs to examine the role of imageability and AoA in 1936 disyllabic words obtained from the English Lexicon Project (ELP; Balota et al., 2007). Cortese and Schock found that the variance in performance of naming latencies and LD reaction times was explained by AoA and imageability beyond word frequency, initial phoneme onset and word length. In a similar study, Juhasz et al. (2015) used the same approach as Cortese and Schock to examine the role of the AoA effect in compound words. They included the baseline model as word frequency, word length and frequency of the morphemes, adding the compound word AoA and morpheme AoA separately and found that the AoA of the compound word, not the morpheme, contributed to LD reaction times beyond the baseline model. Cortese and Schock concluded that semantic effects are larger in longer and more complex words such as disyllabic words than monosyllabic words, as readers take longer to process and compute the pronunciation of the word. This allows semantics to affect the processing of words via interactive activation. In addition, the letter-to-sound correspondence is less predictable in disyllabic words than monosyllabic words, leading to more emphasis on semantics.

Most of these studies were limited to visual LDTs. Turner et al. (1998), however, considered AoA and word frequency in auditory and visual LDTs, and observed independent effects of AoA when controlling for frequency in both conditions. When controlling AoA, word frequency was observed only in the visual, not auditory, LDT (see also Fiebach et al., 2003 who replicated these findings in German). Using auditory LDT, Smith et al. (2006) observed that the AoA effect declined with development but frequency effects were present in children only (but see R.Davies et al., 2017 who observed that with increasing age, the AoA and frequency effects decrease but do not disappear in visual LDT). These studies indicate that the AoA effect is more likely to be found in the links between orthography and phonology to semantics, whereas word frequency is more likely evident in the relationship between orthography and semantics.

The influence of semantic processing and AoA effects on word recognition in opaque languages is well documented but less so in more transparent languages. Burani et al. (2007) examined the role of AoA and word frequency in LDTs in Italian. They observed an interaction between word frequency and

AoA, such that AoA effects were evident only in low-frequency words. Izura and Hernández-Muñoz (2017) concluded that semantic variables such as imageability did not contribute to LD latencies. Nevertheless, they observed that word frequency, word length and AoA contributed to LD latencies, with AoA being larger than word frequency (but see González-Nosti et al., 2014 who observed that semantic variables contributed to LD latencies but were more evident in low-frequency and early-acquired words). This may result from the systematic nature of letter-sound mapping, thus precluding the need for semantics.

The role of AoA in the LDT has been investigated primarily in alphabetic languages, but rarely in Logographic languages. In Mandarin Chinese, Weekes et al. (2008) showed that frequency and AoA contributed to LD latencies (see also Chen et al., 2009; Xu et al., 2020 who replicated these findings). In a more recent and larger-scale study conducted in traditional Chinese, Chang and Lee (2020) used character naming and character decision (i.e. decide whether a Chinese character is a character or not) with Taiwanese students. They observed that early-acquired traditional Chinese characters were responded to more quickly than late-acquired characters. In addition, low-frequency traditional Chinese characters were responded to more slowly than high-frequency characters in these tasks. Chang and Lee also observed that there were non-zero estimates of effects of imageability and semantic ambiguity in character naming and LDT, and that imageability and semantic ambiguity contributed significantly more to character naming than to the character decision task. In addition, Chang and Lee argued that there is more semantic processing in character naming than character decision tasks. The AoA effect was noted to be larger in inconsistent, than consistent words for character naming only. In addition, the AoA effects were found to be larger in character naming than in the LDT. This is the converse of what is shown in alphabetic languages. Chang and Lee concluded that the AoA effect cannot be determined primarily by the arbitrariness of the mappings but also the consolidation and formation of representations. These findings indicate that mapping regularity and access to the semantic representations are both involved in the AoA effect.

It appears that results from the most recent experiments converge on AoA effects contributing to lexical decision reaction times. The results from the LDT seem to support the idea of a frequency-dependent effect, as posited in the representation theory. The size of the AoA effect (45ms on average in Table 6) is equivalent to the size of the frequency effect in lexical decision (50ms). However, the finding from Turner et al. (1998) and Smith et al. (2006) that the AoA effects are observed in an auditory LDT with no corresponding frequency effect is not overtly explained by the representation theory.

Based on the tables, we extracted the means of the AoA effect from LDT and word naming of each study (excluding delayed naming) and compared them using an independent t-test. In support of the representation theory, mapping theory and multiple loci account, AoA effects are smaller in spoken word naming (i.e. 22ms) than LDT (M = 45ms) ($t(64) = 4.23$, $p < .001$, d = 1.09[0.54, 1.64]).

Eye-tracking studies

In contrast to the previous methodologies, which focus on single word recognition and production, eye-tracking is more reflective of natural reading processes. In reading studies, eye-tracking measures usually use whole sentences or paragraphs, which allows the experimenter to assess measures of orthographic, phonological and semantic processing. Depending on how many regions of interest are defined, there are many measures eye-tracking can use to assess lexical retrieval in word and pictorial stimuli. They can be classified into broad categories to reflect the temporal stages of processing: early or late indexes. Early measures including first fixation duration (i.e. duration of initial fixation on a word) and single fixation duration (i.e. a word receives one fixation that allows identification prior to the upcoming word being processed during the first pass of a sentence), can be argued to map onto the initial stages of lexical retrieval such as perceptual features, orthography and phonology. Later measures, which include gaze duration (the sum of fixations on the word prior to the upcoming word in the sentence, analogous to the later stages of lexical retrieval) and total fixation duration (the total duration of processing of words, including re-reading the word) tap into the later stages of lexical retrieval such as semantics and sentential integration (see best practices by Carter & Luke, 2020). However, it is important to consider that when analysing the later eye-tracking measures such as gaze duration, they will also include early measures such as first fixation duration. These measures are not independent from each other[7].

[7] It is important to remember that there is no clear mapping between processing stages and eye movement measures. For instance, on some trials the initial stages of lexical retrieval may occur in first fixation duration, while others might be delayed until gaze duration, whereas semantic processing may be delayed to total time or seen on first fixation duration.

According to the mapping theory, representation theory and multiple loci account, AoA effects should be observed in all measures of eye-tracking. The multiple loci account argues that as each is more likely to involve additional processing stages, there should be larger AoA effects with each additional stage being accessed. However, the representation theory argues that more semantic processing is involved in the later stage than in the earlier stage, hence the AoA effect should be larger in the later stage. Finally, the mapping theory argues that as the later stages of eye-tracking have less regular and systematic mapping between representations (orthography and semantics) than the earlier stages, the AoA effect will be larger there.

As noted in Tables 7 and 8, AoA and frequency effects have been repeatedly demonstrated in text reading in young adults (Juhasz & Rayner, 2003, 2006), in older adults and in alexic patients (Cushman & Johnson, 2011). The AoA effect has been observed in every measure of eye-tracking for early and late stages of reading. However, it is difficult to discern which mechanisms give rise to the AoA effect, as the stimuli for early- and late-acquired words differ from each other in terms of orthography, phonology and meaning. One way to manage this is by using ambiguous words, which only differ in semantics. Juhasz et al. (2011) used ambiguous words (e.g. YARD) that differed in terms of early-acquired meaning (i.e. an area of land next to a home), with the late-acquired meaning (i.e. a distance equal to three feet). Juhasz et al. found that when prior sentential context disambiguated the ambiguous word (e.g. there were weeds everywhere in the **yard** of the office building that Jasper noticed), the AoA effect contributed to the processing of the target word. However, if the sentential context followed the ambiguous word (e.g. Jasper noticed that the **yard** of the office building has weeds everywhere), frequency was more important and AoA was not evident (but see Joseph et al., 2014 who observed that the AoA contributed to word recognition when neutral sentences are used[8]). Juhasz and colleagues (Juhasz & Rayner, 2003, 2006; Juhasz et al., 2011) concluded that the access to various levels of representations (e.g. orthography, phonology, and semantics) in the mental lexicon during reading causes the AoA effect.

[8] The difference between these findings may result from the fact that Juhasz et al. was assessing AoA for known words, whereas Joseph et al. was assessing OoA, for novel nonsense words.

These studies provide information on the AoA effect on fixation duration but have been limited to short and morphologically simple target words. Building on these findings, Juhasz (2018) assessed the role of five variables, including semantic transparency, lexeme meaning dominance, sensory experience rating, imageability familiarity and AoA, on fixation durations of compound words, controlling the frequency and length of the compound word and its lexemes. Juhasz noted that the AoA effect contributed to gaze duration and total reading times when the length of compound word, together with frequency of the compound word and its constituents were controlled. Juhasz (2018) concluded that as the AoA effect is apparent in the later stages of eye-tracking, then the access to the semantic representation during reading causes the AoA effect, at least in compound words.

The AoA effects in eye-tracking were replicated in a further study in which Dirix and Duyck (2017a) asked monolinguals to read an entire novel and recorded their eye movements. The AoA effects were evident in all single, first fixation and gaze durations, together with total reading time (see also Juhasz and Sheridan, 2020 who observed that the AoA effect has an early and enduring effect on every stage of word recognition). Dirix and Duyck (2017b) recorded the eye movements of 19 bilinguals reading half of a novel in their primary language (i.e. Dutch) and the other half in English. Dirix and Duyck (2017b) noted a within-language effect of AoA on second language processing on all timed measures, supporting the mapping theory. However, the mapping theory does not give a specific linguistic level of representation where the AoA effect is observed and the AoA of the second language affects first fixation, gaze duration and total reading time. Furthermore, Dirix and Duyck (2017b) observed that the AoA effect may not be fully independent in L2 (English) and can be influenced by L1 (Dutch) such that long Dutch words with an early AoA facilitated the reading of the English translations, which is in line with the representation theory. Finally, it is noted that the AoA effect increased with each additional stage of eye-tracking such that the AoA effect was largest in the measure that involved semantic processing, aligning with the multiple loci account. These studies indicate that the AoA effect is not a result of task-related artefacts and generalises beyond single word recognition to online reading, also highlighting its important role in lexical processing.

Table 7.

Findings using a regression design in eye-tracking.

Study	Language	Group	N_s	N_i	No of trials	AoA	Frequency	Interaction	Significant predictors	Non-significant predictors
Juhasz and Rayner (2003)	English	YA	40	72	2880	FF ✓[+a] SiF ✓ GD✓[a] TT ✓[a]	FF ✓ SiF ✓ GD✓ TT ✓	FF ✓[a] SiF ✓[+a] GD X TT ✓[+a]	FF: Con[a], Fam SiF: Fam GD Con, Fam, L TT Con, L	FF L SiF Con, L GD **X** TT Fam[+]
Dirix and Duyck (2017a)[b]	English	YA	14	7158	100212	FF 10ms* SiF 14ms* GD 21ms* TT 36ms*	FF 9ms* SiF 13ms* GD 22ms* TT 37ms*	FF X SiF X GD X TT HF: X LF: ✓	FF: L, L * F, L * LP SiF: L*LP GD: L*F TT: L, RoO	FF NHD, LP SiF: L[+], NHD, LP, RoO GD: NHD, LP, RoO TT: LP, NHD
Dirix and Duyck(2017b)[c]	Dutch	YA	19	966/1069	18354/20311	FF ✓[+] SiF n.s. GD ✓ TT ✓	FF ✓ SiF ✓[+] GD ✓ TT ✓	FF X SiF X GD X TT X	FF: La, L2AoA*L1P, L1AoA*La*L SiF: La, L, L2AoA*La, L2*L1P, L1AoA*La*L, La*F*L GD: La, L, L2AoA*La, L2AoA*LD, La*L, F*L, L1AoA*La*L,La*F*L TT:La, L, RoO, L1AoA*La, L2AoA*La, L2AoA*L1P, La*F, La*L, F*L, La*F*L	FF: L[+], L2AoA, L1P, L2P, RoO, LD, La*L1AoA[+], L2AoA*La[+], La*F, La*L, F*L, La*F*L[+] SiF: L1P, L2P, RoO, LD, L1AoA*La, L1AoA*L, L2AoA*LD[+], La*F[+], La*L, GD: L1P, L2P, RoO, LD, L1AoA*La, L1AoA*L, L2AoA*F[+], La*F, TT: L1P, L2P, LD, L1AoA*L, L1AoA*La*L[+]
Juhasz (2018)	English	YA	45	209[d]	9405	FF ✓[+] SiF ✓[+] GD ✓ TT ✓			FF: Fam, SER SiF: Fam GD: Fam TT: Fam	FF: I, ST, LMD SiF I, ST, LMD, SER GD I, ST, LMD, SER TT I, ST, LMD, SER

Note. The language investigated; type of group tested; sample size of participants and items; total number of trials; a significant (✓) at p <.05 or non-significant (X) effect of age of acquisition (AoA) and frequency; interaction between these effects; other variables included in the equation. N_s = Number of subjects; N_i=number of items; YA = young adults; FF = First fixation; SiF =single fixation; GD = gaze duration; TT = total fixation; Con = concreteness; Fam = Familiarity; L = letter length; F = frequency; LP = language proficiency, RoO = Rank of occurrence; NHD = neighbourhood density; La = language; L2AoA = The age of acquisition of words for the second language; L1P = first language proficiency; L1AoA = The age of acquisition of words for the first language; LD = Levenshtein distance; SER = sensory experience rating; I = imageability; ST = semantic transparency; LMD = lexeme meaning dominance; [a] = effect disappears once Word Frequency Guide, analogous to cumulative frequency is included in the equation; [b]= AoA and frequency were a continuous variable, they used a median split to make it categorical and provided the AoA and frequency effect. [c] = Only L1 AoA is mentioned, but there was no main effects of L2AoA, only interactions of L2AoA with word frequency in first fixation, single fixation and a trend for gaze duration, not total fixation duration; [d]= All items were bimorphemic English compound words; [+] .05 < p < .10.

Table 8.

Findings using a factorial design in eye-tracking.

Study	Language	Group	Type of factorial design	N_s	N_i	No of trials	AoA	Frequency	Interaction	Variables controlled
Juhasz and Rayner (2006)	English	YA	Fac	40	72	2880	FF 9* SiF 10* GD 20* TT 26*	FF 9*[a] SiF 10*[a] GD 10*[a] TT 9*[a]	FF X SiF X GD X TT X	P, U, L
Juhasz and Rayner (2006)	English	YA	SF	40	36[b]	1440	FF 15* SiF 17* GD 21* TT 23+	FF 12* SiF 15* GD 14+ TT 20*		L, Con, Fam
Cushman and Johnson (2011)	English	PA	SF[c]	1[d]	27	27	FF N/A SiF N/A GD N/A TT 2236*			BF, Con, Conf, CF, Fam, F, KF, I, L, ON, SL, TF
Cushman and Johnson (2011)	English	OA	SF[c]	6[d]	27	162	FF N/A SiF N/A GD 41* TT 43+			BF, Con, Conf, CF, Fam, F, KF, I, L, ON, SL, TF
Juhasz and Sheridan (2020)	English	YA	SF[c]	47	100	4700	FF 8* SiF 9* GD 9*[a] TT 16*			F, L, Fam, OLD, SR, CP

Note. The language investigated; type of group tested; type of factorial design; sample size of participants and items; total number of trials; the strength of the age-of-acquisition (AoA) and frequency effect (in milliseconds) and whether these effects and interactions were significant at $p < .05$; and variables that were experimentally or statistically controlled. Semi-factorial studies include the sample size of both participants and items for both the AoA- and frequency-manipulated lists separated by a slash, while other studies are stated to be semi-factorial design as they only measured the AoA effect by itself with no other psycholinguistic predictor being included in the same list (see section on statistical analyses and research design of AoA). YA = young adults; OA = older adults; PA = patients with alexia; SF = semi-factorial design, Fac = Factorial design; N_s = Number of subjects; N_i=number of items; FF = First fixation; SiF =single fixation; GD = gaze duration; TT = total fixation; P = number of phoneme; U = understandability; L = length of letters; Con = concreteness; Fam = familiarity; BF = bigram frequency; Conf = confusability; CF = cumulative frequency; Fam = Familiarity; F = frequency; KF = Kucera-Francis frequency, I = imageability; ON = orthographic neighbourhood; SL = syllable length; TF = trigram frequency; OLD = Orthographic Levenstein Distance; SR = sentence rating; CP = cloze probability. [a] = Significant by participant, not items; [b] = Factorial analysis based on items used in the regression study of Juhasz and Rayner (2003); [c]= These studies are seen as semi-factorial designs, as they do not manipulate AoA and an additional psycholinguistic factor (see section on statistical analyses and research design of AoA; [d]= Only paired data is reported.

Progressive demasking

Progressive demasking (PD) procedures entail an alternating word (defined as a signal in the literature) and pattern mask (defined as noise in the literature) being presented to the participant. Through successive display changes, the word gradually emerges from the mask. As soon as the word is identified, participants must type their response (i.e. to ensure the participant correctly recognised the stimuli) as quickly as possible, yielding response times and accuracy measures. The pattern mask is argued to assess visual processing before identification has occurred, thus is a purer measure of the early stages of lexical retrieval (Carreiras et al., 1997) and stretches out the recognition processing, making the task more susceptible to perceptual processing (Dufau et al., 2008). The stimulus presentation for PD is short, thus accuracy ought to depend on whether visual recognition has succeeded. As a result, this task precludes access to semantics and phonological representations but allows perceptual and orthographic processing to occur. It could therefore be predicted that the AoA effect should be present according to the mapping theory and multiple loci account, while the representation theory would predict that there should be no AoA in PD.

Progressive demasking (e.g. Chen et al., 2009; Dent et al., 2007) has been used to investigate the role of the AoA effect during the early stages of word identification. Originally, there was some contention about the role of AoA on PD reaction time and accuracy (Gilhooly & Logie, 1981a, 1981b; Ghyselinck et al., 2004b; Lyons et al., 1978). However, using a factorial design, Chen et al. (2009) observed that the AoA effect was demonstrated using the tachistoscopic presentation of words. It could be concluded that the AoA effects occur at the early stages of word recognition. Furthermore, the evidence seems to indicate that the AoA effect can be found specifically in the connections between orthography and semantic representations. In a PDT, Ploetz and Yates (2016) observed that imageability effects were larger for late-acquired words than early-acquired words. Ploetz and Yates concluded that the imageability effect arose from the semantic feedback to orthographic processing, while orthographic-semantic processes caused the AoA effect.

There is an analogous task that uses pictorial, instead of word stimuli, defined as the visual duration threshold. It is often found that visual duration threshold is lower for early-acquired names than late-acquired names, while controlling for word frequency (B. Chen et al., 2009; Dent et al., 2007). This was extended to include word frequency in the visual duration threshold, highlighting the distinct stages between word frequency and AoA effects. Catling et al. (2008) observed that both AoA and frequency contributed to the visual duration threshold. However, Catling et al. also manipulated the images by adding irrelevant contours, which affects the perceptual processes, while maintaining semantic processes, and observed that the AoA effect was larger under degraded than normal conditions, but the frequency effect remained similar between conditions (see also Preece, 2015, who observed that the AoA, not word frequency, contributed to the visual duration threshold). This indicates that the AoA effect arises at the pre-conceptual stage, while word frequency may contribute to the post-conceptual stage. It is also obvious that the AoA effect is observed in visual duration threshold and PD tasks, as the PD and visual duration threshold use accuracy as outcome variables, they have also been compared with each other (see Tables 9 and 10), and there was no significant difference ($t(6) = 1.13$, $p = .3$, $d = 0.83$ [-1.03, 2.69], $BF_{01} = 0.56$), indicating that the early stages of word recognition and object recognition are being accessed, supporting the multiple loci account and mapping account, that the AoA effect is present in the early stages. This contradicts the representation theory, as the AoA effect is present even when word frequency is controlled for, or is present even when word frequency is being manipulated (e.g. Preece, 2015), which contradicts the notion of a frequency-dependent AoA effect being present in the early stages of word recognition (Belke et al., 2005).

Table 9.

Findings using a regression design in progressive demasking and visual duration threshold tasks.

Study	Task	Presentation	Language	Group	N_s	N_i	No of trials	AoA	Frequency	Interaction	Significant predictors	Non-significant predictors
Gilhooly and Logie (1981b)[a]	PD	V	English	YA	36	100	3600	X	✓		L, Fam, Con	Im, Amb
Gilhooly and Logie (1981b)[b]	PD	V	English	YA	18	100	1800	X	✓		L	Fam, Con, Im, Amb
Gilhooly and Logie (1981b)	PD	A	English	YA	16	100	1600	X	✓[c]		Amb[c]	Con, Fam, L, Im

Note. The type of task; the presentation of stimuli used**;** The language investigated; type of group tested; sample size of participants and items; total number of trials; a significant (✓) at p <.05 or non-significant (X) effect of age of acquisition (AoA) and frequency; interaction between these effects; other variables included in the equation. PD = progressive demasking task; A = auditory; V = visual; N_s = Number of subjects; N_i=number of items; YA = young adults; Amb = ambiguity; Con = concreteness; Fam = Familiarity; L = word length; Im = imagery. [a] = the items were randomly drawn in Experiment 1; [b]= the items were randomly drawn but to reduce any inter-correlations between AoA and other psycholinguistic variables and [c] = Frequency appeared when using a stepwise regression approach, but disappeared when using a simultaneous regression approach, while the converse is demonstrated for the ambiguity variable.

Table 10.

Findings using a factorial design in progressive demasking and visual duration threshold tasks.

Study	Task	Language	Group	Analysis	N_s	N_i	No of trials	AoA	Frequency	Interaction	Variables controlled
Ghyselinck et al. (2004b)	PD	Dutch	YA	Fac	20	192	3840	15%*	10%*	X	L, SL, NHD
Dent et al. (2007)[a]	VDT	English	YA	SF	16	48	768	5.9*			SpF, WF, KF, Fam, NA, P, IA and VC
Dent et al. (2007)[b]	VDT	English	YA	SF	16	48	768	10.6*			SpF, WF, KF, Fam, NA, P, IA and VC
Dent et al. (2007)[c]	VDT	English	YA	SF	16	48	768	4*[d]			SpF, WF, KF, Fam, NA, P, IA and VC
Chen et al. (2009)	PD	Chinese	YA	SF	30	72	2160	5%*			ChF, SC, Con, Fam and H
Chen et al. (2009)	VDT	Chinese	YA	SF	31	72	2232	2*			ChF, SC, Con, Fam and H
Preece (2015)	VDT	English	YA	SF	20/20	68/68	1360/1360	6.07*	0.63		SpF, WF, KF/RAoA, OAoA, Fam, VC, I, PNA, Con, NHD, L, P, SL
Ploetz and Yates (2016)[e]	PD	English	YA	Fac	43	64	2752	7.1%*			F, ON, PN, SL and L

Note. The type of task; the presentation of stimuli used**;** The language investigated; type of group tested; sample size of participants and items; total number of trials; the strength of the age-of-acquisition (AoA) and frequency effect (in milliseconds or percentage correct for progressive demasking) and whether these effects and interactions were significant at p < .05; and variables that were experimentally or statistically controlled. Semi-factorial studies include the sample size of both participants and items for both the AoA- and frequency-manipulated lists separated by a slash, while other studies are stated to be semi-factorial design as they only measured the AoA effect by itself with no other psycholinguistic predictor being included in the same list (see section on statistical analyses and research design of AoA). PD = progressive demasking task; VDT = visual duration threshold task; N_s = Number of subjects; N_i=number of items; YA = young adults; SF = semi-factorial design, Fac = Factorial design; L = word length; SL = syllable length; NHD = neighbourhood density; SpF = spoken frequency; WF = written frequency; KF = Kucera-Francis Frequency; Fam = Familiarity; NA = name agreement; P = number of phonemes; IA = image agreement; VC = visual complexity; ChF = character frequency; SC = Strokes per character; Con = concreteness; H = number of homophones; PNA = picture-name agreement; I = imageability; RAoA = rated age-of-acquisition; OAoA = objective age-of-acquisition; ON = orthographic neighbourhood density; PN = phonological neighbourhood density. [a] = Experiment 1 investigated whether the AoA effect contributed to the VDT ; [b] = Experiment 2 used a degrading manipulated by overlaying the stimuli with a set of outline circles to make it difficult to attain an appropriate structural description of the stimuli. [c]= Experiment 3 reduced contrast and prevent object recognition, thus pictures were presented as high or low contrast images; [d] was significant by subject, approached significance by item; [e] They investigated the role of imageability and AoA in a word identification task and found words with high imageable words were identified more quickly than low imageable words but no interaction, thus is a factorial design

Face recognition

The majority of studies investigating the AoA effect in non-linguistic and naturally occurring stimuli such as faces have used celebrity faces. Originally, models of face naming were developed together with object naming models, and frequency was argued to contribute to both objects and faces. The initial stages in processing objects and faces involve perceptual analysis that produces an internal representation of their visual properties. If the face and object are familiar, a representation of the object and face recognition is activated. However, the processes differ for object and face recognition once semantic processing is involved. The lexical entry for objects (e.g. the lemma *dog*) is activated by several concepts, spreading to semantically related concepts and lemmas so they compete to be chosen, producing semantic interference (Levelt et al., 1999). Afterwards, the name retrieval stage is accessed, where the appropriate phonology is available for the articulatory system. For faces, the stored face recognition unit spreads to the person identity node. These nodes are token markers that activate to define an individual and access semantic information and name representations (Burton et al., 1990). The activation of semantic features activates related identities at the person identity node (e.g. seeing the face of Tony Blair may activate the identity of George Bush). If two known people share semantic information (e.g. both are politicians), they will share semantic nodes at the semantic information units representation. This spreads to the identities at the person identity node level via bidirectional excitatory links. To sum up, semantic activation spreads directly to the lexical level for objects but the person identity node for faces (Valentine et al., 1996). Lexical representations during object naming receives activation from several concepts (i.e. several types of dogs; Levelt et al., 1999), while lexical entries during face naming are only activated by a unique link (e.g. only one Tony Blair). In turn, once proper names are activated, they still receive less activation than common object names.

Furthermore, the AoA measurement differs between object naming and face naming, such that ratings for face naming requires the AoA of a celebrity referring to when an individual is encountered, while for objects it refers to when an individual encounters an object. In addition, it has been argued that the age at which individuals first learn the names of celebrities is later than when they learn the name of objects, thus the AoA for faces may show a qualitative difference to that of knowledge pertaining to new people does not stop but continues to be acquired over one's lifetime, as new individuals become celebrities.

Although the processes underlying object and face recognition are qualitatively different, the AoA effect has repeatedly been shown in face recognition (Lewis, 1999; Lewis et al., 2002; Moore & Valentine, 1998). Importantly, face recognition can therefore assess whether the AoA effect arises at the perceptual or conceptual level, without any influences of orthography. According to the mapping theory, representation theory and multiple loci account, AoA effects should be observed in face recognition. However, according to the multiple loci account and representation account, the AoA effect should be smaller in the early stages of processing than the later stages of processing, as semantic processing is more likely to be necessitated and hence more levels of representation are being accessed in the later stages.

Richards and Ellis (2008) asked participants to make decisions as to whether famous male faces were familiar (i.e. familiarity decision; Experiment 1) and whether they were male or female (i.e. gender decision; Experiment 2). The famous male faces were either presented as whole faces or only the internal features (i.e. eyes, nose and mouth) were exposed. Richards and Ellis observed AoA and rated masculinity effects but the effects were subsumed under an interaction such that rated masculinity only contributed to late-acquired items. The findings were replicated, except for an independent AoA effect in the gender decision task. Following this, Richards and Ellis (2009) included female faces and replicated the findings of AoA and face type impacting familiarity decision and gender decision tasks. Specifically, in the familiarity decision task, Richards and Ellis (2009) observed that the AoA effect was only noted in male, not female, faces. In the gender decision task, Richards and Ellis observed that the AoA effect interacted with face type and gender such that the AoA effect did not contribute to the familiarity decision times when responding to the whole face stimuli, but interacted with gender only when responding to internal features, such that an AoA effect was found in only the internal features in the male face stimuli. However, a reverse AoA effect (i.e. late-acquired female faces were recognised more quickly than early-acquired faces) was observed for the internal features in female faces. Richards and Ellis (2009) concluded that familiarity decision and gender decision have a common basis in semantic representations. However, future research should assess the AoA effect at a perceptual level, by masking the facial features or noise at a perceptual conscious threshold in order to preclude access to semantics (Elsherif et al., 2017).

These findings suggest that the presence of AoA effect in gender decision and face categorisation depends on both the mapping between form and meaning representations (Ellis & Lambon-Ralph, 2000) and is partly semantic in nature (Ghyselinck et al., 2004a; Steyvers & Tenenbaum, 2005). These findings on face recognition cannot be explained by the multiple loci account, representation theory or mapping alone but can be explained by an integrated account of the AoA effect.

Written/Typed picture naming

Most of the studies discussed above make use of oral responses. Similar to spoken picture naming, the AoA effect has been independently observed in written picture naming (Bonin et al., 2001, 2002; Bonin & Meot, 2002). Cuetos et al. (2004) observed that AoA, word frequency and word length contributed to written picture naming in deaf individuals (see also Bonin et al., 2004 who observed that AoA, name agreement and image agreement contributed to both spoken and written naming latencies with the difference being that only letter length contributed to written naming). The AoA effect in written picture naming has been further investigated using event-related potentials to assess the time course of spoken and written AoA. Perret et al. (2014) found that the AoA effect, of 45ms, contributed to both spoken and written picture naming. Using event-related potentials, they observed an AoA effect at around 400 milliseconds after stimuli presentation (i.e., access to the phonological and orthographic forms for spoken and written naming, respectively; Indefrey & Levelt, 2004). Catling and Elsherif (2020) also observed an AoA effect, of 54ms, in written picture naming. In addition, they noted an AoA effect was present in both word-picture verification and written picture naming. Catling and Elsherif wanted to assess if the AoA effect was presented between semantic representations and orthographic representations by subtracting written picture naming from word-picture verification. They observed that an AoA effect was present when subtracting the latencies of object recognition from that of written picture naming. This provides converging evidence that the AoA effect is found within the connection between semantic representations and output form.

In contemporary society, most individuals have moved from handwriting to typing text. The first study to investigate typing the name of a picture was by Snodgrass and Yuditsky (1996). They compared which predictors of typing were similar to spoken picture naming. They found that AoA and name agreement predicted both spoken picture naming and typing. However, typing was affected by concept and image agreement, while frequency contributed to spoken picture naming. This was related to the onset intervals, not inter-keystroke intervals, which assess the processes after the onset. Scaltritti et al. (2016) assessed the variables involved in a typewritten picture naming task at the onset and inter-keystroke intervals. They found that onset latencies were modulated by word frequency, name agreement and AoA, whereas for orthographic variables such as bigram frequency, orthographic similarity, letter length, name agreement and word frequency contributed to within-word intervals.

Written word naming

In contrast to written picture naming, there is a dearth of research exploring the role of AoA in written word naming (i.e. writing down the word presented either visually or auditorily). However, Catling and Elsherif (2020) did observe AoA (12ms) and frequency effects (18ms) in immediate copying (i.e. copying the words presented on the screen). The AoA effect has also been investigated in terms of spelling-to-dictation. Weekes et al. (2006) investigating the AoA effects on 7 to 11-year-old children's spelling. Weekes et al. demonstrated that the AoA effect interacted with orthographic-phonological consistency such that the latter was larger in late-acquired words (see also Weekes et al. 2003, who observed main effects of AoA, not frequency and an interaction of AoA and predictability such that the AoA effect was manifest in unpredictable word spelling). In addition, the AoA, familiarity, character frequency and semantic transparency predictor was also found to contribute to accuracy in Cantonese writing-to-dictation (Su et al., 2022).

Taken together, there are two conclusions: first, the AoA effect occurs at the lexical-semantic level and second, similar to spoken naming, in written naming the AoA effect is larger for pictorial stimuli than word stimuli, as the connections between representations are arbitrary, as opposed to systematic and regular. These findings can be easily incorporated in the representation theory, mapping theory and multiple loci account, as lemmas would be accessed prior to orthographic word form. However, for written word production further investigation is required.

Evaluation of AoA Theories

In this section, experimental evidence providing support for and against the three main AoA theories (mapping theory, representation theory and multiple loci account) is evaluated.

The Multiple Loci Account

According to the multiple loci account, the AoA effect is present in the early stages of reading (i.e. orthographic) and pictorial processing (i.e. perceptual), and semantic and phonological processing. Based on the literature review, supporting evidence has demonstrated that the access to perceptual, phonological and semantic representations during reading and pictorial processing is where the AoA effect comes from. Specifically, AoA effects have been shown in a range of tasks that require lexical access and articulation not necessitating semantic processing (e.g. Barry et al., 2001; Baumann & Ritt, 2018; Elsherif et al., 2019; Gerhand & Barry, 1998, 1999a; Morrison et al., 1992), and in tasks that necessitate access to semantics but not phonology (e.g. Brysbaert et al., 2000; Catling et al., 2021; Catling & Elsherif, 2020; Catling & Johnston, 2006a, 2009; Holmes et al., 2006; Johnston & Barry, 2005; Moore et al., 2004; Morrison & Gibbons, 2006; Palmer & Havelka, 2010; Preece, 2015; Räling et al., 2015, 2016, 2017; Stadthagen-Gonzalez et al., 2004, 2009; Vitkovitch & Tyrrell, 1995 but see Bonin et al., 2006; Chalard & Bonin, 2006).

An additional notable prediction for the multiple loci account is that the AoA effect should be demonstrated in the early and late stages of lexical retrieval (i.e. structural/orthographic and phonological levels of processing). Spataro et al. (2012) used a word-fragment completion task, which involved participants being given fragments of previously studied or unstudied words such as "a_p_e" for apple and told to complete them with the first word that entered their mind. Spataro et al. argued that this task is based primarily on orthographic processes of the studied word and found that repetition priming was larger for late-acquired than for early-acquired items. These findings have been replicated in tasks that assess explicit orthographic representations (e.g. orthographic decision tasks, anagram solutions; Adorni et al., 2013; Gilhooly & Gilhooly, 1979; Stratton et al., 1975 but see Gilhooly & Johnson, 1978). These few studies indicate that the AoA effect indeed occurs in the early stages of word recognition.

However, neuroimaging studies have provided further evidence that the AoA effect emerges in tasks that necessitate access to semantics, but not phonology. An example of such study is Urooj et al.

(2014), who used magnetoencephalography (MEG) to explore the AoA effect on the occipital and left anterior temporal cortex activity during covert object naming (i.e. participants name objects silently, as opposed to ‘out loud’). Urooj et al. observed the AoA effects during covert object naming and noted that the structural properties of an object formed in the occipital cortex is not influenced by AoA. However, there is a fast-forward sweep of activation that results from the occipital and left anterior temporal cortex, causing stronger activation of perceptual-semantic representations for early-acquired objects than for late-acquired objects (but see Perret & Bonin, 2019), indicating that the AoA effect is driven by both phonological and semantic processing.

The Representation theory

According to the representation theory, the AoA effects result from the construction of semantic representations (Brysbaert & Ghyselinck, 2006). Compared to late-acquired words, early-acquired words are at the centre of the semantic network and have stronger connections with other words, thus have richer semantic representations. One prediction from the representation theory is that the frequency and AoA effects will be highly correlated in tasks that do not involve semantic processing, whereas the AoA effect will be larger than the frequency effect in tasks that necessitate access to semantics. Menenti and Burani (2007) found that the AoA effects did not differ between LD latencies and semantic categorisation in both Italian and Dutch. Notably, the AoA coefficients were larger than the frequency coefficients (see also Deyne & Storms, 2007 who observed that the AoA effect was larger for frequency effects in older adults), supporting the representation theory of the AoA effect. In addition, the more semantic processing involved in the task, the larger the AoA effect, which has been shown in reviews by Brysbaert and Ellis (2016), Juhasz (2005) and the current review. The current review calculated the difference in response latencies between early-acquired and late-acquired words to provide a calculation for the AoA effect and observed that the AoA effects are generally larger for picture naming (104ms) than for LDT (45ms) followed by word naming (23ms). This demonstrates that the more semantic processing is necessitated in a task, the larger the AoA effect.

According to Belke et al. (2005), competition arises when a lemma must be selected for a specific concept. Belke et al. provided evidence for larger semantic blocking effects for early-acquired words relative to late-acquired words, indicating that lemmas for early-acquired words may be stronger competitors given a certain concept. As a result, late-acquired words have to compete more strongly against early-acquired competitors. This has been supported by other authors who noted that AoA has a strong effect on the selection of the correct word when several word candidates have been activated and are in competition with each other (Catling et al., 2010; Catling & Johnston, 2005; de Zubicaray et al., 2012; den Hollander et al., 2019; Dent et al., 2008; Karimi & Diaz, 2020; Loftus & Suppes, 1972; Navarrete et al., 2015; Räling et al., 2015; Smith-Spark et al., 2013; Woollams, 2012 but see Dewhurst & Barry, 2006 who noted that the AoA effect did not occur in the Stroop task, which is involved in naming the colour but suppressing the naming of the word). In addition, this pattern of finding is evident in bilingual participants. Using a translation task, Bowers and Kennison (2011) observed that blocking by semantic categories produced longer translation times only in early-acquired words when translating from L1 to L2 but not in late-acquired words.

Furthermore, during vocabulary acquisition, early-acquired concepts serve as the reference point to which late-acquired concepts are compared. In other words, once late-acquired words are learned, they are connected to earlier-acquired concepts and words. For instance, late-acquired words such as 'newborn' would be connected to earlier-acquired words such as baby, child, and so on. As the new word and concept is acquired, this characteristic increases the semantic similarity of earlier-acquired words and concepts (i.e. words established early in development form connections with other words that are semantically similar), whereas a late-acquired word is thus by itself and has not formed as many connections as that of early-acquired words and concepts. As a result, late-acquired words are more semantically distinct than earlier-acquired words. This claim has been supported in recognition memory tasks, as distinctiveness at the lexical level is shown to be positively related to recall and recognition memory, reflected by the lexical predictors (i.e. familiarity, frequency and AoA). Low-frequency words and late-acquired words have more distinct representations than high-frequency and early-acquired words, facilitating recognition memory (Cortese et al., 2010; Cortese et al., 2015; Gullick & Juhasz, 2008). This indicates that late-acquired concepts are semantically distinct. According to the representation theory, both the AoA and frequency effects should be comparable in all tasks that do not necessitate a unique lemma selection. The empirical evidence detailed above gives a different picture. The AoA effect has been found in tasks without frequency effects being observed, such as the VDT, spoken word naming and object recognition. In addition, the AoA and frequency effects are demonstrated in visual LDT but only the AoA effect is noted in the auditory LDT with the same stimuli being used (Turner et al., 1998; see also Smith et al., 2006 who identified the AoA effect over development but found that the frequency effect was present in children only when the same stimuli were used). In the present review, we calculated the correlation, using Pearson's correlation, of the AoA effect and frequency effect in all data published in all naming tasks, excluding delayed naming, or LDT and eye-tracking for factorial designs from 1973 – 2020. We observed that the correlations reported between the frequency and AoA effects were relatively small and negative, amounting to -0.15, when delayed naming was not included, and .70 for lexical decision and .77 for eye-tracking. There are small frequency and AoA effects in word naming, and these variables are larger in LDT, while in picture naming, much larger AoA effects persist in the presence of small, reverse or absent frequency effects. In addition, the negative relationship between frequency and AoA effect, together with the persistence of the AoA effect indicate that the AoA and frequency effects seem to be different in tasks that perhaps do not require access to the mental lexicon such as LDT. It is difficult to see how the representation theory could explain these results.

The representation theory argues that a common semantic system underlies language processing in the participant's first and second languages. If AoA has a semantic basis, the AoA of the primary language would drive performance regardless of the AoA of the second language. In Spanish-English bilinguals, Izura and Ellis (2002) used a picture naming task (Experiment 1) and LDT (Experiment 2) and noted the AoA effect was present for L1 and L2. They also examined whether it was L1 or L2 AoA that produced the AoA effect in L2, by orthogonally manipulating L1 and L2 AoA of their stimuli in Experiment 4, using LDT. They found an AoA effect for the second language, regardless of when the translation equivalent from the first language was learned. The AoA effect seems only to have an impact within each language (see also Izura & Ellis, 2004 who replicated these findings in both translation judgments and LDT), indicating that the AoA effect occurs at the word form level. However, using eye-tracking, Dirix and Duyck (2017b) tested L2 reading in a natural environment and were the first to observe an L1 AoA effect on L2 reading for all timed measures except single fixation duration. The L1 AoA effect arose for L2 long words, and translations of early-acquired words in L1 showed a facilitatory effect on first fixation and gaze duration. Dirix and Duyck concluded that the L1 AoA effect on L2 reading originates from shared semantics across languages, which take time to activate during reading. This is more likely to occur for longer words, as semantics are more likely to be activated, indicating that the representation theory alone is not sufficient to explain these findings (see integrated view of the AoA account).

The Mapping Theory

According to the mapping theory, there is greater structural change for patterns acquired and entered early in a network than those entered and acquired later, producing an advantage for early-trained and acquired patterns and resulting in a gradual decrease in neural plasticity (A.W.Ellis & Lambon Ralph, 2000). An additional prediction from the mapping theory is that the AoA effect should be larger when the relationship between input and output is arbitrary[9]. This prediction was supported in a simulation conducted by P.Monaghan and Ellis (2010), who observed that when input-output mappings are systematic and regular, small effects of AoA were observed, while more inconsistent and unpredictable mappings lead to larger AoA effects. In addition, a large-scale corpus analysis by P. Monaghan et al. (2014) examined the extent to which the mapping was systematic or arbitrary between orthography/phonology and meaning in English vocabulary, specifically onomatopoeia (e.g. woof, roar, meow). They correlated the similarities in terms of phonological mapping and the varying degrees to which phonology maps to semantics. The values for meaning were obtained based on semantic features from WordNet (Miller et al., 1990) and contextual co-occurrence vectors (cf. Latent Semantic Analysis; Landauer et al., 1998). Compared to late-acquired words that have an arbitrary mapping between sound and meaning, P. Monaghan et al. (2014) demonstrated that the mapping between sound and meaning in early-acquired words was more systematic and regular. This systematicity did not differ for nouns and verbs in English (see also L.K. Perry et al. 2015 who replicated the findings behaviourally). The findings emphasise that early-acquired words have non-arbitrary form-meaning mappings that foster word learning and explain that late-acquired words have a more arbitrary nature (i.e. the AoA effect being demonstrated in irregular, not regular, words; see also Monaghan & Ellis, 2002b). These findings

[9] One of the advantages of the mapping theory is that many of the empirical outcomes can be tested via simulations (e.g. Lambon Ralph & Ehsan, 2006, Monaghan & Ellis, 2010, Chang et al., 2019). The precision of the explanation gained from these simulations gives the account more weight than the other theoretical accounts.

underscore that this relationship between sound and meaning is an important property in spoken language learning and contributes to the AoA effect.

According to the mapping theory, patterns trained early in a network cause greater structural changes than later-trained patterns, resulting in an advantage for early-trained patterns and a gradual loss in brain plasticity (Ellis & Lambon Ralph, 2000). One method to test this prediction is learning novel words. In the oral language domain, people acquire novel phonological labels for familiar items such as the *camisa* in Spanish to the word shirt in English and found that early-acquired labels benefitted more than later-acquired labels in being learned, even when cumulative and word frequency is controlled (Catling et al., 2013; Izura et al., 2011; Stewart & Ellis, 2008; Tamminen & Gaskell, 2008). Using eye movements, Joseph et al. (2014) exposed adults to ‘nonce’ words (i.e. nonwords used for one experiment) over a series of 5 days. On Day 1, early-acquired items were introduced, while late-acquired items were presented on Day 2. Joseph et al. found that late-acquired items took longer to read in neutral sentences after training than early-acquired items (see also De Wilde et al., 2020 who noted that early-acquired words in the primary language led to higher vocabulary learning in the secondary language than late-acquired words). These studies show that the general learning mechanism for the AoA effect can be explained by mapping theory.

According to the mapping theory (Monaghan & Ellis, 2002b), AoA effects should be smaller in transparent orthographies than opaque orthographies for word naming. It is argued that the spelling-to-sound correspondence for transparent orthographies is less arbitrary than that of opaque orthographies (R.Davies et al., 2014). The evidence demonstrated in the previous section for visual word naming showed that the AoA effect in languages with opaque orthographies such as English was larger than languages with transparent orthographies. The other two theories (i.e. representation theory and multiple loci account) do not explain why these differences between languages manifest. The different AoA effects observed between languages can be accounted for by the mapping theory only.

However, it is important to note that the mapping theory needs to consider the role of semantics. As high-imageability words have a richer semantic representation, they have a central place in the core of the semantic network (Henry & Kuperman, 2013; Steyvers & Tenenbaum, 2005), facilitating the semantically mediated processing of such words. This makes word processing susceptible to semantic involvement (Raman, 2018; Wilson et al., 2013). However, if the words are abstract or low in imageability or the stimuli contain a plethora of items with low imageability (Raman, 2018; Wilson et al., 2013), the AoA effect is unlikely to be produced (see also Bakhtiar & Weekes, 2015 who observe the same pattern in another transparent language such as Persian). These findings underscore that the mapping theory needs to consider the varying degree of semantic processing needed for items with differing levels of imageability and the importance of orthographic transparency to unravel the AoA effect. Taken together, network plasticity, lexical-semantic competition and the accumulation of each level of processing contribute to the AoA effect during reading, word recognition, production and pictorial processing.

An integrated view of the AoA effect

The representation theory and the multiple loci account focus on the level of representation (i.e. the structural mechanism), whereas the mapping theory and the accumulation hypothesis of the multiple loci account focus on the connections between the level of representations (i.e. functional mechanism). These hypotheses are viewed as being separate, as they assume different mechanisms drive the AoA effect. A recent view is that the AoA effect is the resulting combination of the formation of representations and changing plasticity in the neural network over development (Brysbaert & Ellis, 2016; Catling & Elsherif, 2020; Chang et al., 2019; Chang & Lee, 2020; Cortese et al., 2020; Dirix & Duyck, 2017a; Elsherif & Catling, 2021; Rubin, 1980; Xue et al., 2017).

Within the AoA research, there is a consensus in the general principle of first learned and faster processing. This has been observed throughout this manuscript, especially in the evidence for AoA effects, and evaluation of AoA theories section. In addition, in this manuscript in the evidence for the AoA effect section, we have observed that AoA effects are largest in picture naming tasks followed by LDT and then, word naming. This can be accounted for by the multiple loci account, representation theory and mapping theory. Interestingly, it was noted in the word naming task, that the AoA effect was found to be moderated by the level of orthographic transparency such that in transparent orthographies, the mapping between letter and sound was consistent and regular, thus late-acquired words can benefit from the structure formed by early-acquired words in the neural network, producing a small AoA effect.

However, in opaque orthographies, where the mapping between letter and sound is inconsistent and irregular, late-acquired words cannot benefit from the structured form by early-acquired words in the neural network. The evidence detailed here primarily supports the mapping hypothesis, which is a parsimonious explanation for the AoA effect.

However, as shown within the word naming subsection, orthographic transparency moderates the AoA effect as does imageability. The AoA effect is more likely to be observed in highly imageable words than low-imageable words (e.g. R. Davies et al., 2014; Raman, 2018; Wilson et al., 2013). As the AoA effect is not observed in low-imageable words, this may result from a stronger relationship between letter and sound such that semantic processing does not have time to activate. In contrast to low-imageable words, the visuo/experiential features in high-imageable words are more likely to be activated to a greater extent, leading to more access in semantic processing during reading (Buchanan et al., 2001; Crutch et al., 2009; Pexman et al., 2002), enabling the AoA effects to occur. In contrast, this does not explain the pattern in English, as the AoA effects are larger for low-imageable words than for highly imageable words (Cortese & Schock, 2013). One explanation is that a small AoA effect occurs in highly imageable words, as the AoA effect benefits the relationship between orthographic, phonological and semantic representations are more consistent and regular. In low-imageable words, late-acquired words have a weaker connection between orthography and phonology and so are unable to benefit from the structure produced by early-acquired words, as the mapping between representations is more irregular. There is more time for semantic activation to influence processing, producing a larger AoA effect. To sum up, the different pattern of findings cannot be explained by the arbitrary nature of coding in orthography and phonology or semantics alone. The AoA effect arises as a result of both access to semantic representations (Brysbaert & Ghyselinck, 2006) and less predictable mappings between representations (Lambon Ralph & Ehsan, 2006; Mermillod et al., 2012). Therefore, we can conclude that the integrated account of the AoA effect can explain the mixed findings concerning the relationship between AoA effect and imageability.

In addition, there are several findings that are difficult to reconcile with the mapping theory and representation theory alone but may be explained by an integrated account. Using reading and spelling to dictation tasks, Weekes et al. (2006) observed that AoA interacted with orthographic-phonological consistency in word reading such that the consistency effect was present in late-acquired words only, but phonological-orthographic consistency did not interact with AoA in spelling. The former is easily explained by the mapping theory, as word frequency and AoA share a common neural basis of learning. However, Weekes et al.'s findings result from competition from within the network, which is greater for spelling than reading, hence the lack of interaction between consistency and AoA could be because the number of mappings between input and output is large (for example in picture naming). Also, this could be generalised to typing (Scaltritti et al., 2016), where the response between input and output is also arbitrary, leading to larger AoA effects in these tasks. The results from our literature review and these studies support the integrated view of AoA, showing that lexical processing is gradually shaped by learning during development as a result of more connections and easier access to early-acquired than late-acquired words. The experience of learning influences both representations and mappings in lexical processing.

When studying language, words in different languages can have the same meaning, thus similar or the same conceptual representations and perhaps, the same perceptual representation. These representations are more easily accessed when learned earlier. Early-acquired words can alter a network's weight such that they benefit more from the network plasticity than late-acquired words. Early acquired words also have a more central place in the network for the semantic and perceptual representations. These AoA mechanisms independently and simultaneously affect the speed at which words are processed. As a result, an integrated account can help us to understand the nuances of the AoA effect and unravel its mechanisms.

Recently, using a computational model of reading across development, Chang et al. (2019) included print, sound and meaning of words in their computational model, with training based on children's gradual exposure to language. They also included a computational model of LDT and word naming. Chang et al. observed that AoA effects were stronger for the LDT than for word naming, and noted an interaction between AoA and consistency in word naming such that consistent words produced a smaller AoA effect than inconsistent words, both supporting the mapping theory. Chang et al. showed that words acquired before literacy, where connections involve meaning through sound, demonstrated a weaker AoA effect, whereas an exaggerated AoA effect was observed when words were acquired after literacy, where connections entail meaning via orthography. They concluded that AoA effects arise from both between and within the levels of representation and could be observed simultaneously during incremental learning (Chang et al., 2019).

A majority of AoA studies have focused on alphabetic languages, however, the argument for the integrated hypothesis can be further strengthened by an assessment of its generalisability to different language systems, for example Chinese. The mappings between representations differ between Chinese and alphabetic languages, as Chinese is far more arbitrary in nature. Furthermore, in Chinese, there is a substantial influence of semantics on both lexical decision and character naming tasks. Most studies in Chinese have focused their interpretation on mapping theory, with few studies incorporating representation theory. For instance, Xue et al. (2017) used a semantic priming study and observed larger N100 (i.e. perceptual processing) amplitudes for early-acquired irregular words than late-acquired irregular words in the parietal area and more negative N400 in the frontal and central areas of the brain, indicating that there is an arbitrary mapping between the form of the word (i.e. orthography and phonology and the semantic concept) for early-acquired words. These findings could also be subsumed under a representation theory, as early-acquired words are located at the hub of the semantic network, thus would require less processing, leading to a larger N400, supporting the integrated account.

Further supporting evidence came from Cortese et al. (2020), who analysed data from the English Lexicon Project (ELP) to assess the extent to which the AoA effect results from semantic association with a specific word. Cortese et al. used backward number association (i.e. the number of words that led to the production of the target word in free association) and observed that the AoA effect was reduced but remained a significant contributor to both the latencies of LDT and word naming tasks. This reduction was larger in the LDT than word naming. Cortese et al. concluded that the semantic properties of the AoA reside in the number of backward connections to the word (see also Steyvers & Tenenbaum, 2005 who observed that the early-acquired word is at the central hub of the network, where it spreads its connections to other words) and the remaining variance for the AoA effect results from network plasticity, supporting the integrated account of the AoA effect.

The integrated account is not beyond criticism- some evidence supports the viewpoint that the AoA may occur pre-semantically. Catling and Elsherif (2020) used a collection of methods (i.e. picture-word verification, word-picture verification, spoken picture naming, spoken word naming, written picture naming and written word naming), together with pictorial and word stimuli to indirectly assess if the AoA effect was observed in the links between representations, noting the AoA effect in all tasks. However, in word-picture verification, Catling and Elsherif noted that the AoA effect was larger in the verification than the falsification response. Previous studies (e.g. Bonin et al., 2006; Stadthagen-Gonzalez et al., 2009) have noted that the falsification response taps into the early stages prior to the semantic processes, while verification taps into the perceptual and semantic processes. This suggests that the AoA effect may originate at the perceptual, as opposed to the semantic, level. In addition, Catling and Elsherif subtracted the spoken and written naming from the verification tasks to provide an access score from perceptual/semantics to phonological/orthographic output. They observed that for word stimuli, there was no AoA effect, while for pictorial stimuli, an AoA effect was evident. This shows that when there is an arbitrary relationship between input and output, the AoA effect increases (in line with the accumulation hypothesis), whereas when the input and output relationship is systematic and regular, the AoA effect does not increase. In addition, Catling et al. (2021) used the Recognition without identification (RWI) paradigm, in which people can maintain the ability to recognise a situation as familiar but are unable to recall details of that specific memory. It has been argued that this task reflects the early stages of processing at the perceptual level. Catling et al. observed that the RWI and AoA interacted only for pictorial stimuli, but not for word stimuli, suggesting that the AoA effect originates at the perceptual, not the semantic, loci.

These findings indicate that the AoA effect should pervade the whole of language processing. In addition, early-acquired words benefit not only from multiple connections to other conceptual and perceptual representations but have the opportunity to modify the structure of the neural network, which makes it difficult for late-acquired words to consolidate. Thus, as the relationship between form and meaning becomes more arbitrary, the processing cost increases, leading to a larger AoA effect.

Future research and conclusions

It is a truism that the ability to access the mental lexicon is affected by the AoA and frequency of the word. Compared to late-acquired and low-frequency words, early-acquired words and high-frequency words are easier to retrieve. There are other factors that contribute to lexical retrieval, but rated AoA and word frequency are among the strongest. In addition, if a wide range of values of rated AoA and word frequency and optimal measures of these variables is used, the AoA and frequency effects are found to be similar for tasks that necessitate access to the links between orthographic and phonological representation, while when tasks necessitate further access to additional processes such as semantics and competition between lemmas such as picture naming, the AoA effect is observed to be larger than the word frequency effect. In order to provide a better estimate of the AoA effect and the frequency effect and the causes underlying these effects, it is recommended to use a creative destruction approach (i.e. pre-specifying alternative results by competing hypotheses on a complex set of experimental findings; Tierney et al., 2020, 2021) to work revising theoretical assumptions and, identifying meaningful moderators for further empirical testing' (Delios et al., 2022, p.8).

Based on the current literature review, a majority of studies have been limited to the use of word stimuli. It is important to use other stimuli (such as music) that are more naturalistic and perhaps do not depend on semantics and orthography in order to assess the earlier stages of the AoA (Belfi & Kacirek, 2020) to evaluate whether the AoA effect generalises beyond language. For instance, the AoA effect has been demonstrated in songbird learning, which has been compared to speech and language development (Brainard & Doupe, 2002). In songbird learning, there are three phases: sensory, sensorimotor and the crystallised phase (Brainard & Doupe, 2002). During the sensory phase, a bird is exposed to the tutor's song and forms a template in memory. This song is not crystallised and possesses irregular, idiosyncratic and species-specific properties. During the sensorimotor stage, songs are fine-tuned through practice as the bird learns to match the song of the tutor. In the crystallised period, birds mature and produce a species-specific song but are most often unable to learn new songs due to computational demands. However, if they do learn a new song, it takes them longer to master it (Lipkind et al., 2017; see also the mapping theory). These studies highlight that the AoA effect is a property of learning in both humans and songbirds. By investigating the AoA effect in music in humans, one can assess the earlier stages of the AoA effect such as the auditory perceptual level prior to semantics being involved, which allows researchers to evaluate whether the AoA effect occurs pre-conceptually. Finally, it is known that music can be used to relieve anxiety and depression in Alzheimer's disease (Moreira et al., 2018), while larger than expected AoA effect is found to be a predictor of mild cognitive impairment (Catling et al., 2013). Investigations into how these two components interact could help devise a useful tool to predict mild cognitive impairment and improve the quality of life of those suffering from degenerative diseases.

Beyond music, the AoA effect has also been investigated in motor learning. To our knowledge, only one study has shown that AoA affects motor learning and retention. Magill (1976) asked participants to learn three positions on a manual lever in a serial order. Magill found that position 1 took less time to learn and resulted in fewer errors than positions 2 and 3. In addition, they observed that reaction time and errors were smaller for position 2 than position 3. This indicates that the AoA effect is not due to the use of verbal material but to the nature of the series being learned (Neumann & Ammons, 1957). Similar to word learning studies conducted on the order of acquisition, it could be suggested that motor learning works in a similar manner. If this is the case, and following the integrated account, early-acquired motor tasks would enter the motor domain, which has plasticity, leading to rich, stable and established sensorimotor representations. The connections between sensory and motor representations are modified by these early-acquired motor tasks, engendering the network to lose plasticity. As a result, when a novel and late-acquired motor task enters, it becomes more difficult to consolidate, allowing early-acquired items to be processed and produced more quickly than late-acquired items. In addition, early-acquired motor tasks would have more connections to other motor tasks than late-acquired motor tasks, which have fewer connections and perhaps may be retained less. By investigating the AoA effect in motor learning in humans, one can assess whether early learning occurs using more sensorimotor processing relative to late learning. This would not only enable us to generalise the AoA effects beyond language to motor and music but also to consider that "when naturalistic learning conditions prevail, it is the absence rather than the presence of age (or order) of acquisition effects that would represent a challenge to computational models that have learning as a central tenet "(P. Monaghan & Ellis, 2010), thus providing us an understanding that the AoA effect is a property emerging out of any learning system. A further outcome of this review is that it would appear sensible to provide novel AoA ratings for each sample and take into account the specific characteristics of the sample (e.g. older and younger populations). It is unclear whether the lack of group differences results from the sampling or that the estimates have been assessed using young, as opposed to older, adults.

Although word frequency and objective AoA are perceived to be more objective measures, there is a subjective aspect underlying both measures. Older adults and young adults populations have different learning histories and may produce cohort effects, thus encountering words differently from younger populations. For instance, words related to products and technology may be rated as early-acquired and high-frequency by young adults but later-acquired and low-frequency by elderly populations. It is therefore necessary to obtain AoA and frequency ratings from elderly participants to better predict the performance of neurotypical adults and patients with acquired conditions such as Alzheimer's disease (Cuetos et al., 2012; Deyne & Storms, 2007) or use several word frequencies to assess the corpus best used for the participant sample tested. To assess the influence of word frequency or AoA between young and elderly populations, we need to manipulate both recent and dated measures, not one or the other, otherwise our understanding of the effect of AoA and word frequency on lexical retrieval will be limited in the long-run. By evaluating and refining norming studies, researchers would be able to obtain more precise measures of AoA and word frequency to assess the lifespan and development of lexical retrieval.

Finally, the AoA effect has been investigated across languages and neurotypes (e.g. dyslexia and stuttering), using megastudies, cross-task comparisons and a wide range of tasks. Based on the evidence, we conclude that the role of the AoA effect on lexical retrieval is defined by the influence of individual learning experiences on the gradual development of representations, and the mappings between representations. This highlights that the AoA effect pertains to information about learning throughout development. Put simply, the AoA effect could thus be a principle of general learning in terms of cognition that, in future, will contribute to a life span theory of lexical retrieval.

References

Adorni, R., Manfredi, M., & Proverbio, A. M. (2013). Since when or how often? Dissociating the roles of age of acquisition (AoA) and lexical frequency in early visual word processing. *Brain and Language*, *124*(1), 132–141. https://doi.org/10.1016/j.bandl.2012.11.005

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Álvarez, B., & Cuetos, F. (2007). Objective age of acquisition norms for a set of 328 words in Spanish. *Behavior Research Methods*, *39*(3), 377–383. https://doi.org/10.3758/BF03193006

Anderson, J. D. (2008). Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *Journal of fluency disorders*, *33*(2), 135-155. https://doi.org/10.1016/j.jfludis.2008.04.001

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic bulletin & review*, *4*(4), 439-461. https://doi.org/10.3758/BF03214334

Arnon, I., McCauley, S. M., & Christiansen, M. H. (2017). Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and Language*, *92*, 265–280. https://doi.org/10.1016/j.jml.2016.07.004

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database*. [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Bakhtiar, M., Nilipour, R., & Weekes, B. S. (2013). Predictors of timed picture naming in Persian. *Behavior Research Methods*, *45*(3), 834–841. https://doi.org/10.3758/s13428-012-0298-6

Bakhtiar, M., & Weekes, B. (2015). Lexico-semantic effects on word naming in Persian: Does age of acquisition have an effect? *Memory & Cognition*, *43*(2), 298–313. https://doi.org/10.3758/s13421-014-0472-4

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual Word Recognition of Single-Syllable Words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Bangalore, S., Robson, H., & Astell, A. J. (2022). Standardizing norms for 180 coloured Snodgrass and Vanderwart pictures in Kannada language. *PloS one*, *17*(4), e0266359. https://doi.org/10.1371/journal.pone.0266359

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 424–434. https://doi.org/10.3758/BF03195471

Barrow, C. W., Björkenstam, K. N., & Strömbergsson, S. (2019). Subjective ratings of age-of-acquisition: Exploring issues of validity and rater reliability. *Journal of Child Language*, *46*(2), 199–213. https://doi.org/10.1017/S0305000918000363

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of Acquisition, Word Frequency, and the Locus of Repetition Priming of Picture Naming. *Journal of Memory and Language*, *44*(3), 350–375. https://doi.org/10.1006/jmla.2000.2743

Barry, C., Johnston, R. A., & Wood, R. F. (2006). Effects of age of acquisition, age, and repetition priming on object naming. *Visual Cognition*, *13*(7–8), 911–927. https://doi.org/10.1080/13506280544000101

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart Pictures: Effects of Age of Acquisition, Frequency, and Name Agreement. *The Quarterly Journal of Experimental Psychology Section A*, *50*(3), 560–585. https://doi.org/10.1080/783663595

Bauer, P. J., Burch, M. M., Scholin, S. E., & Güler, O. E. (2007). Using Cue Words to Investigate the Distribution of Autobiographical Memories in Childhood. *Psychological Science*, *18*(10), 910–916. https://doi.org/10.1111/j.1467-9280.2007.01999.x

Baumann, A., & Ritt, N. (2018). The basic reproductive ratio as a link between acquisition and change in phonotactics. *Cognition*, *176*, 174–183. https://doi.org/10.1016/j.cognition.2018.03.005

Belfi, A. M., & Kacirek, K. (2021). The famous melodies stimulus set. *Behavior Research Methods, 53*(1), 34-48. https://doi.org/10.3758/s13428-020-01411-6

Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition*, *96*(2), B45–B54. https://doi.org/10.1016/j.cognition.2004.11.006

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways?. *Quarterly Journal of Experimental Psychology*, *68*(7), 1268-1294. https://doi.org/10.1080/17470218.2014.978877

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language*, *50*(4), 456–476. https://doi.org/10.1016/j.jml.2004.02.001

Bonin, P., Chalard, M., Méot, A., & Barry, C. (2006). Are age-of-acquisition effects on object naming due simply to differences in object recognition? Comments on Levelt (2002). *Memory & Cognition*, *34*(5), 1172–1182. https://doi.org/10.3758/BF03193263

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2001). Age-of-acquisition and word frequency in the lexical decision task: Further evidence from the French language. *Cahiers de Psychologie Cognitive - Current Psychology of Cognition*, *20*(6), 401–443.

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*(1), 89–114. https://doi.org/10.1348/000712602162463

Bonin, P., Fayol, M., & Chalard, M. (2001). Age of Acquisition and Word Frequency in Written Picture Naming. *The Quarterly Journal of Experimental Psychology Section A*, *54*(2), 469–489. https://doi.org/10.1080/713755968

Bonin, P., & Méot, A. (2002). Writing to dictation in real time in adults: What are the determinants of written latencies. In *Advances in psychology research* (pp.139-165). New York: Nova Science Publishers.

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways?. *Quarterly Journal of Experimental Psychology*, *68*(7), 1268-1294. https://doi.org/10.1080/17470218.2014.978877

Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 158–167. https://doi.org/10.3758/BF03195507

Bonin, P., Perret, C., Méot, A., Ferrand, L., & Mermillod, M. (2008). Psycholinguistic norms and face naming times for photographs of celebrities in French. *Behavior Research Methods*, *40*(1), 137–146. https://doi.org/10.3758/BRM.40.1.137

Bowers, J. M., & Kennison, S. M. (2011). The Role of Age of Acquisition in Bilingual Word Translation: Evidence from Spanish-English Bilinguals. *Journal of Psycholinguistic Research*, *40*(4), 275–289. https://doi.org/10.1007/s10936-011-9169-z

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature*, *417*(6886), 351-358. https://doi.org/10.1038/417351a

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, *15*(3), 208–216. https://doi.org/10.3758/BF03197718

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., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior research methods*, *49*(4), 1520-1523. https://doi.org/10.3758/s13428-016-0811-4

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology*, *64*(3), 545–559. https://doi.org/10.1080/17470218.2010.503374

Brysbaert, M., & Ellis, A. W. (2016). Aphasia and age of acquisition: Are early-learned words more resilient? *Aphasiology*, *30*(11), 1240–1263. https://doi.org/10.1080/02687038.2015.1106439

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition*, *13*(7–8), 992–1011. https://doi.org/10.1080/13506280544000165

Brysbaert, M., Lagrou, E., & Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, *20*(3), 530–548. https://doi.org/10.1017/S1366728916000353

Brysbaert, M., Lange, M., & Wijnendaele, I. V. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology*, *12*(1), 65–85. https://doi.org/10.1080/095414400382208

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, *44*(4), 991–997. https://doi.org/10.3758/s13428-012-0190-4

Brysbaert, M., Wijnendaele, I. V., & Deyne, S. D. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, *104*(2), 215–226. https://doi.org/10.1016/S0001-6918(00)00021-4

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, *8*(3), 531-544. https://doi.org/10.3758/BF03196189

Burani, C., Arduino, L. S., & Barca, L. (2007). Frequency, not age of acquisition, affects Italian word naming. *European Journal of Cognitive Psychology*, *19*(6), 828–866. https://doi.org/10.1080/09541440600847946

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, *81(*3), 361–380. https://doi.org/10.1111/j.2044-8295.1990.tb02367

Butler, B., & Hains, S. (1979). Individual differences in word recognition latency. *Memory & Cognition*, *7*(2), 68–76. https://doi.org/10.3758/BF03197587

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *23*, 857–871. https://doi.org/10.1037//0278-7393.23.4.857

Carroll, J. B., & White, M. N. (1973). Word Frequency and Age of Acquisition as Determiners of Picture-Naming Latency. *Quarterly Journal of Experimental Psychology*, *25*(1), 85–95. https://doi.org/10.1080/14640747308400325

Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, *155*, 49–62. https://doi.org/10.1016/j.ijpsycho.2020.05.010

Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Catling, J. C., Pymont, C., Johnston, R. A., Elsherif, M. M., Clark, R., & Kendall, E. (2021). Age of acquisition effects in recognition without identification tasks. *Memory,* 1-13. https://doi.org/10.1080/09658211.2021.1931695

Catling, J. C., Dent, K., Preece, E., & Johnston, R. (2013). Age-of-Acquisition Effects in Novel Picture Naming: A Laboratory Analogue. *Quarterly Journal of Experimental Psychology*, *66*(9), 1756–1763. https://doi.org/10.1080/17470218.2013.764903

Catling, J. C., Dent, K., & Williamson, S. (2008). Age of acquisition, not word frequency affects object recognition: Evidence from the effects of visual degradation. *Acta Psychologica*, *129*(1), 130–137. https://doi.org/10.1016/j.actpsy.2008.05.005

Catling, J. C., & Elsherif, M. M. (2020). The hunt for the age of acquisition effect: It's in the links! *Acta Psychologica*, *209*, 103138. https://doi.org/10.1016/j.actpsy.2020.103138

Catling, J. C., & Johnston, R. A. (2006a). Age of acquisition effects on an object-name verification task. *British Journal of Psychology*, *97*(1), 1–18. https://doi.org/10.1348/000712605X53515

Catling, J. C., & Johnston, R. A. (2006b). The effects of age of acquisition on an object classification task. *Visual Cognition*, *13*(7–8), 968–980. https://doi.org/10.1080/13506280544000138

Catling, J. C., & Johnston, R. A. (2006c). Effects of age of acquisition and priming on picture naming. *Quarterly Journal of Experimental Psychology*, *59*(8), 1443–1453. https://doi.org/10.1080/17470210500214291

Catling, J. C., & Johnston, R. A. (2009). The varying effects of age of acquisition. *Quarterly Journal of Experimental Psychology*, *62*(1), 50–62. https://doi.org/10.1080/17470210701814352

Chalard, M., & Bonin, P. (2006). Age-of-acquisition effects in picture naming: Are they structural and/or semantic in nature? *Visual Cognition*, *13*(7–8), 864–883. https://doi.org/10.1080/13506280544000084

Chalard, M., Bonin, P., Méot, A., Boyer, B., & Fayol, M. (2003). Objective age-of-acquisition (AoA) norms for a set of 230 object names in French: Relationships with psycholinguistic variables, the English data from Morrison et al. (1997), and naming latencies. *European Journal of Cognitive Psychology*, *15*(2), 209–245. https://doi.org/10.1080/09541440244000076

Chang, Y.-N., & Lee, C.-Y. (2020). Age of acquisition effects on traditional Chinese character naming and lexical decision. *Psychonomic Bulletin & Review*, 1–8. https://doi.org/10.3758/s13423-020-01787-8

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

Chedid, G., Wilson, M. A., Bedetti, C., Rey, A. E., Vallet, G. T., & Brambati, S. M. (2019). Norms of conceptual familiarity for 3,596 French nouns and their contribution in lexical decision. *Behavior research methods, 51*(5), 2238-2247. https://doi.org/10.3758/s13428-018-1106-8

Chen, B., Dent, K., You, W., & Wu, G. (2009). Age of acquisition affects early orthographic processing during Chinese character recognition. *Acta Psychologica*, *130*(3), 196–203. https://doi.org/10.1016/j.actpsy.2008.12.004

Chen, B. G., Zhou, H. X., Dunlap, S., & Perfetti, C. A. (2007). Age of acquisition effects in reading Chinese: Evidence in favour of the arbitrary mapping hypothesis. *British Journal of Psychology*, *98*(3), 499–516. https://doi.org/10.1348/000712606X165484

Colombo, L., & Burani, C. (2002). The Influence of Age of Acquisition, Root Frequency, and Context Availability in Processing Nouns and Verbs. *Brain and Language*, *81*(1–3), 398–411. https://doi.org/10.1006/brln.2001.2533

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. *Quarterly Journal of Experimental Psychology*, *60*(8), 1072–1082. https://doi.org/10.1080/17470210701315467

Cortese, M. J., Khanna, M. M., & Hacker, S. D. (2010). Recognition memory for 2,578 monosyllabic words. *Memory*, *18*, 595-609. https://doi.org/10.1080/09658211.2010.493892

Cortese, M. J., McCarty, D. P., & Schock, J. (2015). A mega recognition memory study of 2897 disyllabic words. *Quarterly Journal of Experimental Psychology*, *68*, 1489-1501. https://doi.org/10.1080/17470218.2014.945096

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *Quarterly Journal of Experimental Psychology*, *66*(5), 946–972. https://doi.org/10.1080/17470218.2012.722660

Cortese, M. J., Toppi, S., Khanna, M. M., & Santo, J. B. (2020). AoA effects in reading aloud and lexical decision: Locating the (semantic) locus in terms of the number of backward semantic associations. *Quarterly Journal of Experimental Psychology*, 174702182094030. https://doi.org/10.1177/1747021820940302

Cortese, M. J., Yates, M., Schock, J., & Vilks, L. (2018). Examining word processing via a megastudy of conditional reading aloud. *Quarterly Journal of Experimental Psychology*, *71*(11), 2295–2313. https://doi.org/10.1177/1747021817741269

Crutch, S. J., Connell, S., & Warrington, E. K. (2009). The different representational frameworks underpinning abstract and concrete knowledge: Evidence from odd-one-out judgements. *Quarterly Journal of Experimental Psychology, 62*(7), 1377-1390. https://doi.org/10.1080%2F17470210802483834

Cuetos, F., & Alija, M. (2003). Normative data and naming times for action pictures. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 168–177. https://doi.org/10.3758/BF03195508

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology*, *18*(3), 415–436. https://doi.org/10.1080/13594320500165896

Cuetos, F., Herrera, E., & Ellis, A. W. (2010). Impaired word recognition in Alzheimer's disease: The role of age of acquisition. *Neuropsychologia*, *48*(11), 3329–3334. https://doi.org/10.1016/j.neuropsychologia.2010.07.017

Cuetos, F., Samartino, T., & Ellis, A. W. (2012). Age acquisition norms from elderly Spanish people: Characteristics and the prediction of word recognition performance in Alzheimer's disease. *Psicológica*, *33*(1), 19.

Cushman, C. L., & Johnson, R. L. (2011). Age-of-Acquisition Effects in Pure Alexia. *Quarterly Journal of Experimental Psychology*, *64*(9), 1726–1742. https://doi.org/10.1080/17470218.2011.556255

Cutler, A. (1981). Making up materials is a confounded nuisance, or: Will we able to run any psycholinguistic experiments at all in 1990?. *Cognition, 10*, 65-70. doi: 10.1016/0010-0277(81)90026-3.

Davies, R. A. I., Arnell, R., Birchenough, J. M. H., Grimmond, D., & Houlson, S. (2017). Reading through the life span: Individual differences in psycholinguistic effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1298–1338. https://doi.org/10.1037/xlm0000366

Davies, R., Wilson, M., Cuetos, F., & Burani, C. (2014). Reading in Spanish and Italian: Effects of Age of Acquisition in Transparent Orthographies? *Quarterly Journal of Experimental Psychology*, *67*(9), 1808–1825. https://doi.org/10.1080/17470218.2013.872155

Davies, R., Barbón, A., & Cuetos, F. (2013). Lexical and semantic age-of-acquisition effects on word naming in Spanish. *Memory & Cognition*, *41*(2), 297–311. https://doi.org/10.3758/s13421-012-0263-8

Delios, A., Clemente, E. G., Wu, T., Tan, H., Wang, Y., Gordon, M., Viganola, D., Chen, Z., Dreber Johannesson, M., Pfeiffer, T., Generalizability Tests Forecasting Collaboration, & Uhlmann, E. L. (2022). Examining the generalizability of research findings from archival data. *Proceedings of the National Academy of Sciences, 119*(30), e2120377119. https://doi.org/10.1073/pnas.2120377119

De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cognitive and Behavioral Neurology*, *21*(4), 227-235. https://doi.org/ 10.1097/WNN.0b013e318190d162

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English Through Out-of-School Exposure: How Do Word-Related Variables and Proficiency Influence Receptive Vocabulary Learning? *Language Learning*, *70*(2), 349–381. https://doi.org/10.1111/lang.12380

Dell'acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, *32*(4), 588–615. https://doi.org/10.3758/BF03200832

Dent, K., Catling, J. C., & Johnston, R. A. (2007). Age of acquisition affects object recognition: Evidence from visual duration thresholds. *Acta Psychologica*, *125*(3), 301–318. https://doi.org/10.1016/j.actpsy.2006.08.004

Dewhurst, S. A., & Barry, C. (2006). Dissociating word frequency and age of acquisition: The klein effect revived (and reversed). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 919–924. https://doi.org/10.1037/0278-7393.32.4.919

Deyne, S. D., & Storms, G. (2007). Age-of-acquisition differences in young and older adults affect latencies in lexical decision and semantic categorization. *Acta Psychologica*, *124*(3), 274–295. https://doi.org/10.1016/j.actpsy.2006.03.007

Dirix, N., & Duyck, W. (2017a). An eye movement corpus study of the age-of-acquisition effect. *Psychonomic Bulletin & Review*, *24*(6), 1915–1921. https://doi.org/10.3758/s13423-017-1233-8

Dirix, N., & Duyck, W. (2017b). The first- and second-language age of acquisition effect in first- and second-language book reading. *Journal of Memory and Language*, *97*, 103–120. https://doi.org/10.1016/j.jml.2017.07.012

Dufau, S., Stevens, M., & Grainger, J. (2008). Windows executable software for the progressive demasking task. *Behavior Research Methods, 40*, 33–37. http://dx.doi.org/10.3758/BRM.40.1.33

Ellis, A.W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1103–1123. https://doi.org/10.1037/0278-7393.26.5.1103

Ellis, A.W., & Monaghan, J. (2002). Reply to Strain, Patterson, and Seidenberg (2002). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 215–220. https://doi.org/10.1037/0278-7393.28.1.215

Ellis, A.W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 515–523. https://doi.org/10.1037/0278-7393.24.2.515

Ellis, A.W., Burani, C., Izura, C., Bromiley, A., & Venneri, A. (2006). Traces of vocabulary acquisition in the brain: Evidence from covert object naming. *NeuroImage*, *33*(3), 958–968. https://doi.org/10.1016/j.neuroimage.2006.07.040

Elsherif, M. M., Catling, J. C., & Frisson, S. (2020). Two words as one: A multi-naming investigation of the age-of-acquisition effect in compound-word processing. *Memory & Cognition*, *48*(4), 511–525. https://doi.org/10.3758/s13421-019-00986-6

Elsherif, M. M., & Catling, J. C. (2021). Age of acquisition effects on the decomposition of compound words. *Journal of Cognitive Psychology*, 1-14. https://doi.org/10.1080/20445911.2021.2013246

Elsherif, M. M., Saban, M. I., & Rotshtein, P. (2017). The perceptual saliency of fearful eyes and smiles: A signal detection study. *PLOS ONE*, *12*(3), 1–30. https://doi.org/10.1371/journal.pone.0173199.

Elsherif, M. M., Wheeldon, L. R., & Frisson, S. (2022a). Phonological precision for word recognition in skilled readers. *Quarterly Journal of Experimental Psychology, 17470218211046350*, 1–20. https://doi.org/10.1177/17470218211046350.

Elsherif, M. M., Frisson, S., & Wheeldon, L.R. (2022b). Orthographic precision for word naming in skilled readers. *Language, Cognition and Neuroscience.* https://doi.org/10.1080/23273798.2022.2108091

Ernestus, M., & Cutler, A. (2015). BALDEY: A database of auditory lexical decisions. *The Quarterly Journal of Experimental Psychology*, *68*, 1469–1488. doi:10.1080/17470218.2014.984730

Fasquel, A., Brunellière, A., & Knutsen, D. (2022). A modified procedure for naming 332 pictures and collecting norms: Using tangram pictures in psycholinguistic studies. *Behavior Research Methods*, 1-23. https://doi.org/10.3758/s13428-022-01871-y

Fiebach, C. J., Friederici, A. D., Müller, K., von Cramon, D. Y., & Hernandez, A. E. (2003). Distinct brain representations for early and late learned words. *NeuroImage*, *19*(4), 1627–1637. https://doi.org/10.1016/S1053-8119(03)00227-1

Forster, K. I. (2000). The potential for experimenter bias effects in word recognition experiments. *Memory & cognition*, *28*(7), 1109-1115. https://doi.org/10.3758/BF03211812

Francis, W., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Gertel, V. H., Karimi, H., Dennis, N. A., Neely, K. A., & Diaz, M. T. (2020). Lexical frequency affects functional activation and accuracy in picture naming among older and younger adults. *Psychology and Aging*, *35*(4), 536–552. https://doi.org/10.1037/pag0000454

Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 267–283. https://doi.org/10.1037/0278-7393.24.2.267

Gerhand, S., & Barry, C. (1999a). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory & Cognition*, *27*(4), 592–602. https://doi.org/10.3758/BF03211553

Gerhand, S., & Barry, C. (1999b). Age-of-acquisition and frequency effects in speeded word naming. *Cognition*, *73*(2), B27–B36. https://doi.org/10.1016/S0010-0277(99)00052-9

Ghyselinck, M., Custers, R., & Brysbaert, M. (2004a). The Effect of Age of Acquisition in Visual Word Processing: Further Evidence for the Semantic Hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 550–554. https://doi.org/10.1037/0278-7393.30.2.550

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004b). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica*, *115*(1), 43–67. https://doi.org/10.1016/j.actpsy.2003.11.002

Gilhooly, K. J., & Gilhooly, M. L. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory & Cognition*, *7*(3), 214–223. https://doi.org/10.3758/BF03197541

Gilhooly, K. J., & Johnson, C. E. (1978). Effects of Solution Word Attributes on Anagram Difficulty: A Regression Analysis. *Quarterly Journal of Experimental Psychology*, *30*(1), 57–70. https://doi.org/10.1080/14640747808400654

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior research methods & instrumentation*, *12*(4), 395-427. https://doi.org/10.3758/BF03201693

Gilhooly, K. J., & Logie, R. H. (1981a). Word age-of-acquisition and visual recognition thresholds. *Current Psychology*, *1*(3–4), 215–225. https://doi.org/10.1007/BF03186732

Gilhooly, K. J., & Logie, R. H. (1981b). Word age-of-acquisition, reading latencies and auditory recognition. *Current Psychology*, *1*(3–4), 251–262. https://doi.org/10.1007/BF03186735

Gilhooly, K. J., & Logie, R. H. (1982). Word age-of-acquisition and lexical decision making. *Acta Psychologica*, *50*(1), 21–34. https://doi.org/10.1016/0001-6918(82)90048-8

González-Nosti, M., Barbón, A., Rodríguez-Ferreiro, J., & Cuetos, F. (2014). Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words. *Behavior Research Methods*, *46*(2), 517–525. https://doi.org/10.3758/s13428-013-0383-5

Göz, İ., Tekcan, A. İ., & Erciyes, A. A. (2017). Subjective age-of-acquisition norms for 600 Turkish words from four age groups. *Behavior Research Methods*, *49*(5), 1736–1746. https://doi.org/10.3758/s13428-016-0817-y

Grainger, J., & Segui, J. (1990). Neighbourhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & psychophysics*, *47*(2), 191-198.

Gullick, M. M., & Juhasz, B. J. (2008). Age of acquisition's effect on memory for semantically associated word pairs. *Quarterly Journal of Experimental Psychology*, *61*(8), 1177-1185. https://doi.org/10.1080/17470210802013391

Havelka, J., & Tomita, I. (2006). Age of acquisition in naming Japanese words. *Visual Cognition*, *13*(7–8), 981–991. https://doi.org/10.1080/13506280544000156

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior research methods*, *50*(3), 1166-1186. https://doi.org/10.3758/s13428-017-0935-1

Henry, R., & Kuperman, V. (2013). Semantic growth of morphological families in English. *Psihologija*, *46*(4), 479–495. https://doi.org/10.2298/PSI1304479H

Herdağdelen, A., & Marelli, M. (2017). Social Media and Language Processing: How Facebook and Twitter Provide the Best Frequency Estimates for Studying Word Recognition. *Cognitive Science*, *41*(4), 976–995. https://doi.org/10.1111/cogs.12392

Hernandez, A. E., & Fiebach, C. J. (2006). The brain bases of reading late learned words: Evidence from functional MRI. *Visual Cognition*, *13*(7–8), 1027–1043. https://doi.org/10.1080/13506280544000183

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, *133*(4), 638–650. https://doi.org/10.1037/0033-2909.133.4.638

Hirsh, K. W., & Funnell, E. (1995). Those old, familiar things: Age of acquisition, familiarity and lexical access in progressive aphasia. *Journal of neurolinguistics*, *9*(1), 23-32. https://doi.org/10.1016/0911-6044(95)00003-8

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition*, *13*(7–8), 884–910. https://doi.org/10.1080/13506280544000093

Hsiao, Y., & Nation, K. (2018). Semantic diversity, frequency and the development of lexical quality in children's word reading. *Journal of Memory and Language*, *103*, 114–126. https://doi.org/10.1016/j.jml.2018.08.005

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101-144. https://doi.org/10.1016/j.cognition.2002.06.001

Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica*, *23*(2).

Izura, C., & Ellis, A. W. (2004). Age of acquisition effects in translation judgement tasks. *Journal of Memory and Language*, *50*(2), 165-181. https://doi.org/10.1016/j.jml.2003.09.004

Izura, C., & Hernández-Muñoz, N. (2017). The Role of Semantics in Spanish Word Recognition: An Insight from Lexical Decision and Categorization Tasks. *Open Linguistics*, *3*(1), 500–515. https://doi.org/10.1515/opli-2017-0025

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language*, *64*(1), 32–58. https://doi.org/10.1016/j.jml.2010.09.002

Johnston, R. A., & Barry, C. (2005). Age of acquisition effects in the semantic processing of pictures. *Memory & Cognition*, *33*(5), 905–912. https://doi.org/10.3758/BF03193084

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition*, *13*(7–8), 789–845. https://doi.org/10.1080/13506280544000066

Johnston, R. A., Dent, K., Humphreys, G. W., & Barry, C. (2010). British-English norms and naming times for a set of 539 pictures: The role of age of acquisition. *Behavior Research Methods*, *42*(2), 461–469. https://doi.org/10.3758/BRM.42.2.461

Joseph, H. S. S. L., Wonnacott, E., Forbes, P., & Nation, K. (2014). Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition*, *133*(1), 238–248. https://doi.org/10.1016/j.cognition.2014.06.015

Juhasz, B. J. (2005). Age-of-Acquisition Effects in Word and Picture Identification. *Psychological Bulletin*, *131*(5), 684–712. https://doi.org/10.1037/0033-2909.131.5.684

Juhasz, B. J. (2018). Experience with compound words influences their processing: An eye movement investigation with English compound words. *Quarterly Journal of Experimental Psychology*, *71*(1), 103–112. https://doi.org/10.1080/17470218.2016.1253756

Juhasz, B. J., Gullick, M. M., & Shesler, L. W. (2011). The Effects of Age-of-Acquisition on Ambiguity Resolution: Evidence from Eye Movements. *Journal of Eye Movement Research*, *4*(1), 1–14. https://doi.org/10.16910/jemr.4.1.4

Juhasz, B. J., Lai, Y.-H., & Woodcock, M. L. (2015). A database of 629 English compound words: Ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience. *Behavior Research Methods*, *47*(4), 1004–1019. https://doi.org/10.3758/s13428-014-0523-6

Juhasz, B. J., & Rayner, K. (2003). Investigating the Effects of a Set of Intercorrelated Variables on Eye Fixation Durations in Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1312–1318. https://doi.org/10.1037/0278-7393.29.6.1312

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition*, *13*(7–8), 846–863. https://doi.org/10.1080/13506280544000075

Juhasz, B. J., & Sheridan, H. (2020). The time course of age-of-acquisition effects on eye movements during reading: Evidence from survival analyses. *Memory & Cognition*, *48*(1), 83–95. https://doi.org/10.3758/s13421-019-00963-z

Karimi, H., & Diaz, M. (2020). When phonological neighborhood density both facilitates and impedes: Age of acquisition and name agreement interact with phonological neighborhood during word production. *Memory & Cognition*. https://doi.org/10.3758/s13421-020-01042-4

Khwaileh, T., Body, R., & Herbert, R. (2014). A Normative Database and Determinants of Lexical Retrieval for 186 Arabic Nouns: Effects of Psycholinguistic and Morpho-Syntactic Variables on Naming Latency. *Journal of Psycholinguistic Research*, *43*(6), 749–769. https://doi.org/10.1007/s10936-013-9277-z

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Kuperman, V. (2013). Accentuate the positive: Semantic access in English compounds. *Frontiers in Psychology*, *4*(203), 1–10. doi:10.3389/fpsyg.2013.00203

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

Laganaro, M., & Perret, C. (2011). Comparing Electrophysiological Correlates of Word Production in Immediate and Delayed Naming Through the Analysis of Word Age of Acquisition Effects. *Brain Topography*, *24*(1), 19–29. https://doi.org/10.1007/s10548-010-0162-x

Laganaro, M., Valente, A., & Perret, C. (2012). Time course of word production in fast and slow speakers: A high density ERP topographic study. *NeuroImage*, *59*(4), 3881–3888. https://doi.org/10.1016/j.neuroimage.2011.10.082

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*(7–8), 928–948. https://doi.org/10.1080/13506280544000110

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

Laws, K. R., Leeson, V. C., & Gale, T. M. (2002). The Effect of 'Masking' on Picture Naming. *Cortex*, *38*(2), 137–147. https://doi.org/10.1016/S0010-9452(08)70646-4

Lewis, M. B. (1999). Age of acquisition in face categorisation: Is there an instance-based account? *Cognition*, *71*(1), B23–B39. https://doi.org/10.1016/S0010-0277(99)00020-7

Lewis, M. B., Chadwick, A. J., & Ellis, H. D. (2002). Exploring a neural-network account of age-of-acquisition effects using repetition priming of faces. *Memory & Cognition*, *30*(8), 1228–1237. https://doi.org/10.3758/BF03213405

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

Levelt, W. J. M., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75. https://doi.org/10.1017/S0140525X99001776

Lipkind, D., Zai, A. T., Hanuschkin, A., Marcus, G. F., Tchernichovski, O., & Hahnloser, R. H. (2017). Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nature communications*, *8*(1), 1-11. https://doi.org/10.1038/s41467-017-01436-0

Li, L., Yang, Y., Song, M., Fang, S., Zhang, M., Chen, Q., & Cai, Q. (2022). CCLOWW: A grade-level Chinese children's lexicon of written words. *Behavior Research Methods*, 1-16. https://doi.org/10.3758/s13428-022-01890-9

Lima, D., Pinto, R., & Albuquerque, P. B. (2021). Recognition and naming test of the Portuguese population for national and international celebrities. *Behavior Research Methods, 53*(6), 2326-2337. https://doi.org/10.3758/s13428-021-01572-y

Liu, Y., Hao, M., Li, P., & Shu, H. (2011). Timed Picture Naming Norms for Mandarin Chinese. *PLoS ONE*, *6*(1), e16505. https://doi.org/10.1371/journal.pone.0016505

Liu, Y., Hao, M., Shu, H., Tan, L. H., & Weekes, B. S. (2008). Age-of-acquisition effects on oral reading in Chinese. *Psychonomic Bulletin & Review*, *15*(2), 344–350. https://doi.org/10.3758/PBR.15.2.344

Lotto, L., Surian, L., & Job, R. (2010). Objective age of acquisition for 223 Italian words: Norms and effects on picture naming speed. *Behavior Research Methods*, *42*(1), 126–133. https://doi.org/10.3758/BRM.42.1.126

Łuniewska, M., Haman, E., Armon-Lotem, S., Etenkowski, B., Southwood, F., Anđelković, D., Blom, E., Boerma, T., Chiat, S., de Abreu, P. E., Gagarina, N., Gavarró, A., Håkansson, G., Hickey, T., de López, K. J., Marinis, T., Popović, M., Thordardottir, E., Blažienė, A., … Ünal-Logacev, Ö. (2016). Ratings of age of acquisition of 299 words across 25 languages: Is there a cross-linguistic order of words? *Behavior Research Methods*, *48*(3), 1154–1177. https://doi.org/10.3758/s13428-015-0636-6

Łuniewska, M., Wodniecka, Z., Miller, C. A., Smolík, F., Butcher, M., Chondrogianni, V., Hreich, E. K., Messarra, C., A. Razak, R., Treffers-Daller, J., Yap, N. T., Abboud, L., Talebi, A., Gureghian, M., Tuller, L., & Haman, E. (2019). Age of acquisition of 299 words in seven languages: American English, Czech, Gaelic, Lebanese Arabic, Malay, Persian and Western Armenian. *PLOS ONE*, *14*(8), e0220611. https://doi.org/10.1371/journal.pone.0220611

Lyons, A. W., Teer, P., & Rubenstein, H. (1978). Age-at-acquisition and word recognition. *Journal of Psycholinguistic Research*, *7*(3), 179–187. https://doi.org/10.1007/BF01067041

Macmillan, M. B., Neath, I., & Surprenant, A. M. (2021). Re-assessing age of acquisition effects in recognition, free recall, and serial recall. *Memory & Cognition*, *49*(5), 939-954. https://doi.org/10.3758/s13421-021-01137-6

Magill, R. A. (1976). Order of acquisition of the parts of a serial-motor task. *Research Quarterly. American Alliance for Health, Physical Education and Recreation*, *47*(1), 134-139. https://doi.org/10.1080/10671315.1976.10615350

Mainz, N., Shao, Z., Brysbaert, M., & Meyer, A. S. (2017). Vocabulary Knowledge Predicts Lexical Processing: Evidence from a Group of Participants with Diverse Educational Backgrounds. *Frontiers in Psychology*, *8*, 1164. https://doi.org/10.3389/fpsyg.2017.01164

Mandera, P., Keuleers, E., Wodniecka, Z., & Brysbaert, M. (2015). Subtlex-pl: Subtitle-based word frequency estimates for Polish. *Behavior Research Methods*, *47*(2), 471–483. https://doi.org/10.3758/s13428-014-0489-4

Marful, A., Díez-Álamo, A. M., Plaza-Navas, S., & Fernandez, A. (2018). A normative study for photographs of celebrities in Spain. *PLoS ONE, 13*(5), 1–20. https://doi.org/10.1371/journal.pone.0197554

McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the Word Frequency Effect: The Neglected Role of Distributional Information in Lexical Processing. *Language and Speech*, *44*(3), 295–322. https://doi.org/10.1177/00238309010440030101

Meehl, P. E. (1970). Nuisance variables and the ex post facto design. In M. Radner & S. Winokur (Eds.), *Minnesota Studies in The Philosophy of Science* (Vol. 4, pp. 373-402). Minneapolis, MN: University of Minnesota Press.

Menenti, L., & Burani, C. (2007). What causes the effect of age of acquisition in lexical processing? *Quarterly Journal of Experimental Psychology*, *60*(5), 652–660. https://doi.org/10.1080/17470210601100126

Meschyan, G., & Hernandez, A. (2002). Age of acquisition and word frequency: Determinants of object-naming speed and accuracy. *Memory & Cognition*, *30*(2), 262–269. https://doi.org/10.3758/BF03195287

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. I*nternational journal of lexicography*, *3*(4), 235-244. https://doi.org/10.1093/ijl/3.4.235

Monaghan, J., & Ellis, A. W. (2002a). Age of acquisition and the completeness of phonological representations. *Reading and Writing*, *15*(7/8), 759–788. https://doi.org/10.1023/A:1020958722472

Monaghan, J., & Ellis, A. W. (2002b). What exactly interacts with spelling—Sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 183–206. https://doi.org/10.1037/0278-7393.28.1.183

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*(4), 506–525. https://doi.org/10.1016/j.jml.2010.08.003

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130299. https://doi.org/10.1098/rstb.2013.0299

Montefinese, M., Vinson, D., Vigliocco, G., & Ambrosini, E. (2019). Italian Age of Acquisition Norms for a Large Set of Words (ItAoA). *Frontiers in Psychology*, *10*, 278. https://doi.org/10.3389/fpsyg.2019.00278

Moore, V., Smith-Spark, J., & Valentine, T. (2004). The effects of age of acquisition on object perception. *European Journal of Cognitive Psychology*, *16*(3), 417–439. https://doi.org/10.1080/09541440340000097

Moore, V., & Valentine, T. (1998). The Effect of Age of Acquisition on Speed and Accuracy of Naming Famous Faces. *The Quarterly Journal of Experimental Psychology Section A*, *51*(3), 485–513. https://doi.org/10.1080/713755779

Moore, V., Valentine, T., & Turner, J. (1999). Age-of-acquisition and cumulative frequency have independent effects. *Cognition*, *72*(3), 305–309. https://doi.org/10.1016/S0010-0277(99)00042-6

Moreira, S. V., Justi, F. R. D. R., & Moreira, M. (2018). Can musical intervention improve memory in Alzheimer's patients? Evidence from a systematic review. *Dementia & neuropsychologia*, *12*(2), 133-142. https://doi.org/10.1590/1980-57642018dn12-020005

Morrison, C. M. (2003). Interpret with Caution: Multicollinearity in Multiple Regression of Cognitive Data. *Perceptual and Motor Skills*, *97*(1), 80–82. https://doi.org/10.2466/pms.2003.97.1.80

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of Acquisition Norms for a Large Set of Object Names and Their Relation to Adult Estimates and Other Variables. *The Quarterly Journal of Experimental Psychology Section A*, *50*(3), 528–559. https://doi.org/10.1080/027249897392017

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 116–133. https://doi.org/10.1037/0278-7393.21.1.116

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, *91*(2), 167–180. https://doi.org/10.1348/000712600161763

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*(6), 705–714. https://doi.org/10.3758/BF03202720

Morrison, C. M., & Gibbons, Z. C. (2006). Lexical determinants of semantic processing speed. *Visual Cognition*, *13*(7–8), 949–967. https://doi.org/10.1080/13506280544000129

Morrison, C. M., Hirsh, K. W., Chappell, T., & Ellis, A. W. (2002). Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology*, *14*(4), 435–459. https://doi.org/10.1080/09541440143000159

Morrison, C. M., Hirsh, K. W., & Duggan, G. B. (2003). Age of Acquisition, Ageing, and Verb Production: Normative and Experimental Data. *The Quarterly Journal of Experimental Psychology Section A*, *56*(4), 1–26. https://doi.org/10.1080/02724980244000594

Nagy, W., Anderson, R. C., Schommer, M., Scott, J. A., & Stallman, A. C. (1989). Morphological Families in the Internal Lexicon. *Reading Research Quarterly*, *24*(3), 262. https://doi.org/10.2307/747770

Navarrete, E., Scaltritti, M., Mulatti, C., & Peressotti, F. (2013). Age-of-acquisition effects in delayed picture-naming tasks. *Psychonomic Bulletin & Review*, *20*(1), 148–153. https://doi.org/10.3758/s13423-012-0310-2

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics, 28*, 661–677. doi:10.1017/S014271640707035X

Neumann, E., & Ammons, R. B. (1957). Acquisition and long-term retention of a simple serial perceptual-motor skill. *Journal of experimental Psychology*, *53*(3), 159. https://doi.org/10.1037/h0041719

Newcombe, N. S. (2003). Some controls control too much. *Child Development*, *74*(4), 1050-1052. https://doi.org/10.1111/1467-8624.00588

Nishimoto, T., Miyawaki, K., Ueda, T., Une, Y., & Takahashi, M. (2005). Japanese normative set of 359 pictures. *Behavior Research Methods*, *37*(3), 398–416. https://doi.org/10.3758/BF03192709

Palmer, S. D., & Havelka, J. (2010). Age of acquisition effects in vocabulary learning. *Acta Psychologica*, *135*(3), 310–315. https://doi.org/10.1016/j.actpsy.2010.08.002

Parsons, S., Azevedo, F., Elsherif, M. M., Guay, S., Shahim, O. H., Govaart, G. H., Norris, E., O'Mahony, A., Parker, A. J., Todorovic, A., Pennington, C. R., Garcia-Pelegrin, E., Lazić, A., Robertson, O., Middleton, S. L., Valentini, B., McCuaig, J., Baker, B. J., Collins, E., … Aczel, B. (2022). A community-sourced glossary of open scholarship terms. *Nature Human Behavior*, *6*, 312–318. https://doi.org/10.1038/s41562-021-01269-4

Pathman, T., Larkina, M., Burch, M. M., & Bauer, P. J. (2013). Young Children's Memory for the Times of Personal Past Events. *Journal of Cognition and Development*, *14*(1), 120–140. https://doi.org/10.1080/15248372.2011.641185

Perez, M., & Navalon, C. (2005). Objective-AoA norms for 175 names in Spanish: Relationships with other psycholinguistic variables, estimated AoA, and data from other languages. *European Journal of Cognitive Psychology*, *17*(2), 179–206. https://doi.org/10.1080/09541440340000547

Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1

Perret, C., Bonin, P., & Laganaro, M. (2014). Exploring the multiple-level hypothesis of AoA effects in spoken and written object naming using a topographic ERP analysis. *Brain and Language*, *135*, 20–31. https://doi.org/10.1016/j.bandl.2014.04.006

Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and Its Relation to Lexical Category and Age of Acquisition. *PLOS ONE*, *10*(9), e0137147. https://doi.org/10.1371/journal.pone.0137147

Pexman, P. M., Lupker, S. J., & Hino, Y. (2002). The impact of feedback semantics in visual word recognition: Number-of-features effects in lexical decision and naming tasks. *Psychonomic bulletin & review*, *9*(3), 542-549. https://doi.org/10.3758/BF03196311

Pind, J., Jónsdóttir, H., Gissurardóttir, H., & Jónsson, F. (2000). Icelandic Norms for the Snodgrass and Vanderwart (1980) Pictures: Name and Image Agreement, Familiarity, and Age of Acquisition. *Scandinavian Journal of Psychology*, *41*(1), 41–48. https://doi.org/10.1111/1467-9450.00169

Pind, J., & Tryggvadóttir, H. B. (2002). Determinants of picture naming times in Icelandic. *Scand J Psychol*, *43*, 221–226.

Ploetz, D. M., & Yates, M. (2016). Age of acquisition and imageability: A cross-task comparison: Age of Acquisition and Imageability. *Journal of Research in Reading*, *39*(1), 37–49. https://doi.org/10.1111/1467-9817.12040

Preece, E. (2015). *The Loci of Age of Acquisition and Word Frequency Effects: Evidence From Contemporary Experimental Paradigms and Eye-Tracking* (Doctoral dissertation, University of Worcester).

Räling, R., Hanne, S., Schröder, A., Keßler, C., & Wartenburger, I. (2017). Judging the Animacy of Words: The Influence of Typicality and Age of Acquisition in a Semantic Decision Task.

Quarterly Journal of Experimental Psychology, *70*(10), 2094–2104. https://doi.org/10.1080/17470218.2016.1223704

Räling, R., Holzgrefe-Lang, J., Schröder, A., & Wartenburger, I. (2015). On the influence of typicality and age of acquisition on semantic processing: Diverging evidence from behavioural and ERP responses. *Neuropsychologia*, *75*, 186–200. https://doi.org/10.1016/j.neuropsychologia.2015.05.031

Räling, R., Schröder, A., & Wartenburger, I. (2016). The origins of age of acquisition and typicality effects: Semantic processing in aphasia and the ageing brain. *Neuropsychologia*, *86*, 80–92. https://doi.org/10.1016/j.neuropsychologia.2016.04.019

Raman, I. (2006). On the age-of-acquisition effects in word naming and orthographic transparency: Mapping specific or universal? *Visual Cognition*, *13*(7–8), 1044–1053. https://doi.org/10.1080/13506280500153200

Raman, I. (2011). The role of age of acquisition in picture and word naming in dyslexic adults. *British Journal of Psychology*, *102*(3), 328-339. https://doi.org/10.1348/000712610X522572

Raman, I. (2018). The role of context on age of acquisition effect: Strategic control in word naming in Turkish. In D. Buğa & M. Coşgun Ögeyik (Eds.), *Psycholinguistics and Cognition in Language Processing*. IGI Global. https://doi.org/10.4018/978-1-5225-4009-0

Ramanujan, K., & Weekes, B. S. (2020). Predictors of lexical retrieval in Hindi–English bilingual speakers. *Bilingualism: Language and Cognition*, *23*(2), 265–273. https://doi.org/10.1017/S1366728918001177

Richards, R. M., & Ellis, A. W. (2008). Mechanisms of Identity and Gender Decisions to Faces: Who Rocked in 1986? *Perception*, *37*(11), 1700–1719. https://doi.org/10.1068/p6023

Richards, R. M., & Ellis, A. W. (2009). Identity, gender, and the role of age of acquisition in face processing. *Psicológica*, *30*, 155–179.

Rochford, G., & Williams, M. (1962a). I The relationship between nominal dysphasia and the acquisition of vocabulary in childhood. *Journal of Neurology, Neurosurgery & Psychiatry*, *25*(3), 222–227. https://doi.org/10.1136/jnnp.25.3.222

Rochford, G., & Williams, M. (1962b). II Experimental production of naming disorders in normal people. *Journal of Neurology, Neurosurgery & Psychiatry*, *25*(3), 228–233. https://doi.org/10.1136/jnnp.25.3.228

Roodenrys, S., Hulme, C., Alban, J., Ellis, A. W., & Brown, G. D. A. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition*, *22*(6), 695–701. https://doi.org/10.3758/BF03209254

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in methods and practices in psychological science*, *1*(1), 27-42. https://doi.org/10.1177/2515245917745629

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. https://doi.org/10.1016/S0022-5371(80)90415-6

Scaltritti, M., Arfé, B., Torrance, M., & Peressotti, F. (2016). Typing pictures: Linguistic processing cascades into finger movements. *Cognition*, *156*, 16–29. https://doi.org/10.1016/j.cognition.2016.07.006

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270-1281. https://doi.org/10.3758/BF03201199

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, *27*(5), 499–520. https://doi.org/10.1016/0749-596X(88)90022-8

Schwitter, V., Boyer, B., Méot, A., Bonin, P., & Laganaro, M. (2004). French normative data and naming times for action pictures. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 564–576. https://doi.org/10.3758/BF03195603

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological review*, *96*(4), 491-528. 10.1037/0033-295X.106.3.491

Sereno, S. C., & O'Donnell, P. J. (2009). Participant and Word Gender in Age of Acquisition Effects: The Role of Gender Socialization. *Sex Roles*, *61*(7–8), 510–518. https://doi.org/10.1007/s11199-009-9649-x

Severens, E., Lommel, S. V., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

Shibahara, N., & Kondo, T. (2002). Variables Affecting Naming Latency for Japanese Kanji: A Re-Analysis of Yamazaki, *et al.* (1997). *Perceptual and Motor Skills*, *95*(3), 741–745. https://doi.org/10.2466/pms.2002.95.3.741

Smith, M. A., Cottrell, G. W., & Anderson, K. L. (2001). The Early Word Catches the Weights. *Proceedings of the 2000 Conference*, 52–58.

Smith, P. T., Turner, J. E., Brown, P. A., & Henry, L. A. (2006). The Distinct Contributions of Age of Acquisition and Word Frequency in Auditory Word Perception. *Quarterly Journal of Experimental Psychology*, *59*(12), 2121–2134. https://doi.org/10.1080/17470210500438387

Smith-Spark, J. H., & Moore, V. (2009). The representation and processing of familiar faces in dyslexia: Differences in age of acquisition effects. *Dyslexia*, *15*(2), 129-146. https://doi.org/10.1002/dys.365

Smith-Spark, J. H., Moore, V., & Valentine, T. (2012). Long-term age of acquisition effects in famous name processing. *Acta Psychologica*, *139*(1), 202–211. https://doi.org/10.1016/j.actpsy.2011.10.002

Smith-Spark, J. H., Moore, V., & Valentine, T. (2013). Determinants of famous name processing speed: Age of acquisition versus semantic connectedness. *Acta Psychologica*, *142*(2), 230–237. https://doi.org/10.1016/j.actpsy.2012.11.016

Smolík, F., & Filip, M. (2022). Corpus-based age of word acquisition: Does it support the validity of adult age-of-acquisition ratings?. *Plos one*, *17*(5), e0268504. https://doi.org/10.1371/journal.pone.0268504

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods*, *Instruments, & Computers*, *28*(4), 516–536. https://doi.org/10.3758/BF03200540

Spataro, P., Longobardi, E., Saraulli, D., & Rossi-Arnaud, C. (2013). Interactive Effects of Age-of-Acquisition and Repetition Priming in the Lexical Decision Task: A Multiple-Loci Account. *Experimental Psychology*, *60*(4), 235–242. https://doi.org/10.1027/1618-3169/a000192

Spataro, P., Mulligan, N. W., Longobardi, E., & Rossi-Arnaud, C. (2012). Effects of Age-of-Acquisition in the Word-Fragment Completion Task: Evidence for an Orthographic Locus in Implicit Memory. *Experimental Psychology*, *59*(1), 22–29. https://doi.org/10.1027/1618-3169/a000121

Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods*, *14*(2), 287-305. https://doi.org/10.1177/1094428110369842

Stadthagen-Gonzalez, H., Bowers, J. S., & Damian, M. F. (2004). Age-of-acquisition effects in visual word recognition: Evidence from expert vocabularies. *Cognition*, *93*(1), B11–B26. https://doi.org/10.1016/j.cognition.2003.10.009

Stadthagen-Gonzalez, H., Damian, M. F., Pérez, M. A., Bowers, J. S., & Marín, J. (2009). Name–picture verification as a control measure for object naming: A task analysis and norms for a large set of pictures. *Quarterly Journal of Experimental Psychology*, *62*(8), 1581–1597. https://doi.org/10.1080/17470210802511139

Stewart, N., & Ellis, A. W. (2008). Order of acquisition in learning perceptual categories: A laboratory analogue of the age-of-acquisition effect? *Psychonomic Bulletin & Review*, *15*(1), 70–74. https://doi.org/10.3758/PBR.15.1.70

Steyvers, M., & Tenenbaum, J. B. (2005). The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, *29*(1), 41–78. https://doi.org/10.1207/s15516709cog2901_3

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1140–1154. https://doi.org/10.1037/0278-7393.21.5.1140

Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling—Sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 207–214. https://doi.org/10.1037/0278-7393.28.1.207

Stratton, R. P., Jacobus, K. A., & Brinley, B. (1975). Age-of-acquisition, imagery, familiarity and meaningfulness norms for 543 words. *Behavior Research Methods & Instrumentation*, *7*(1), 1–6. https://doi.org/10.3758/BF03201278

Su, I. F., Yum, Y. N., & Lau, D. K. Y. (2022). Hong Kong Chinese character psycholinguistic norms: ratings of 4376 single Chinese characters on semantic radical transparency, age-of-acquisition, familiarity, imageability, and concreteness. *Behavior Research Methods*, 1-20. https://doi.org/10.3758/s13428-022-01928-y

Tamminen, J., & Gaskell, M. G. (2008). Short Article: Newly Learned Spoken Words Show Long-Term Lexical Competition Effects. *Quarterly Journal of Experimental Psychology*, *61*(3), 361–371. https://doi.org/10.1080/17470210701634545

Taylor, E., Nation, K., & Hsiao, Y. (2022). Context Availability and Sentence Availability Ratings for 3,000 English Words and their Association with Lexical Processing. *Journal of Cognition, 5*(1), 1–10. https://doi.org/10.5334/joc.211

Tierney, W., Hardy, J., Ebersole, C. R., Viganola, D., Clemente, E. G., Gordon, M., Hoogeveen, S., Haaf, J., Dreber, A., Johannesson, M., Pfeiffer, T., Huang, J. L., Vaughn, L. A., DeMarree, K., Igou, E. R., Chapman, H., Gantman, A., Vanaman, M., Wylie, J., … Uhlmann, E. L. (2021). A creative destruction approach to replication: Implicit work and sex morality across cultures. *Journal of Experimental Social Psychology, 93*, 1–18. https://doi.org/10.1016/j.jesp.2020.104060.

Tierney, W., Hardy, J. H., Ebersole, C. R., Leavitt, K., Viganola, D., Clemente, E. G., Gordon, M., Dreber, A., Johannesson, M., Pfeiffer, T., & Uhlmann, E. L. (2020). Creative destruction in science. *Organizational Behavior and Human Decision Processes, 161*, 291–309. https://doi.org/10.1016/j.obhdp.2020.07.002.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, *124*(2), 107-136. https://doi.org/10.1037//0096-3445.124.2.107.

Turner, J. E., Valentine, T., & Ellis, A. W. (1998). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. *Memory & Cognition*, *26*(6), 1282–1291. https://doi.org/10.3758/BF03201200

Urooj, U., Cornelissen, P. L., Simpson, M. I. G., Wheat, K. L., Woods, W., Barca, L., & Ellis, A. W. (2014). Interactions between visual and semantic processing during object recognition revealed by modulatory effects of age of acquisition. *NeuroImage*, *87*, 252–264. https://doi.org/10.1016/j.neuroimage.2013.10.058

Valentine, T., Brennen, T., & Brédart, S. (1996). *The cognitive psychology of proper names: On the importance of being Ernest.* London: Routledge.

Valente, A., Burki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: A trial by trial multiple regression analysis from stimulus onset to response. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00390

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), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vitkovitch, M., & Tyrrell, L. (1995). Sources of Disagreement in Object Naming. *The Quarterly Journal of Experimental Psychology Section A*, *48*(4), 822–848. https://doi.org/10.1080/14640749508401419

Wang, J., & Chen, B. (2020). A database of Chinese-English bilingual speakers: Ratings of the age of acquisition and familiarity. *Frontiers in Psychology, 11*, 554785. https://doi.org/10.3389/fpsyg.2020.554785

Weekes, B., Davies, R., Parris, B., & Robinson, G. (2003). Age of acquisition effects on spelling in surface dysgraphia. *Aphasiology*, *17*(6–7), 563–584. https://doi.org/10.1080/02687030344000030

Weekes, B.S., Chan, A. H. D., & Tan, L. H. (2008). Effects of age of acquisition on brain activation during Chinese character recognition. *Neuropsychologia*, *46*(7), 2086–2090. https://doi.org/10.1016/j.neuropsychologia.2008.01.020
Weekes, B.S., Castles, A. E., & Davies, R. A. (2006). Effects of Consistency and Age of Acquisition on Reading and Spelling among Developing Readers. *Reading and Writing*, *19*(2), 133–169. https://doi.org/10.1007/s11145-005-2032-6
Weekes, B.S., Shu, H., Hao, M., Liu, Y., & Tan, L. H. (2007). Predictors of timed picture naming in Chinese. *Behavior Research Methods*, *39*(2), 335–342. https://doi.org/10.3758/BF03193165
Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PloS one*, *11*(3), e0152719. https://doi.org/10.1371/journal.pone.0152719
Whaley, C. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*, 143–154. https://doi.org/10.1016/S0022-5371(78)90110-X
Wilson, M. A., Cuetos, F., Davies, R., & Burani, C. (2013). Revisiting age-of-acquisition effects in Spanish visual word recognition: The role of item imageability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1842–1859. https://doi.org/10.1037/a0033090
Wilson, M. A., Ellis, A. W., & Burani, C. (2012). Age-of-acquisition affects word naming in Italian only when stress is irregular. *Acta Psychologica*, *139*(3), 417–424. https://doi.org/10.1016/j.actpsy.2011.12.012
Wolna, A., Łuniewska, M., Haman, E., & Wodniecka, Z. (2022). Polish norms for a set of colored drawings of 168 objects and 146 actions with predictors of naming performance. Behavior Research Methods, 1-27. https://doi.org/10.3758/s13428-022-01923-3
Wysocki, A., Lawson, K. M., & Rhemtulla, M. (2020, October 13). Statistical Control Requires Causal Justification. *PsyArXiv* https://doi.org/10.31234/osf.io/j9vw4
Xu, X., Li, J., & Guo, S. (2020). Age of acquisition ratings for 19,716 simplified Chinese words. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01455-8
Xue, J., Liu, T., Marmolejo-Ramos, F., & Pei, X. (2017). Age of Acquisition Effects on Word Processing for Chinese Native Learners' English: ERP Evidence for the Arbitrary Mapping Hypothesis. *Frontiers in Psychology*, *8*, 818. https://doi.org/10.3389/fpsyg.2017.00818
Yamada, J., Takashima, H., & Yamazaki, M. (1998). Effect of Ease-of-Acquisition on Naming Latency for Japanese Kanji: A Reanalysis of Yamazaki, *et al.'s* (1997) Data. *Psychological Reports*, *83*(3), 991–1002. https://doi.org/10.2466/pr0.1998.83.3.991
Yamazaki, M., Ellis, A. W., Morrison, C. M., & Ralph, M. A. L. (1997). Two age of acquisition effects in the reading of Japanese Kanji. *British Journal of Psychology*, *88*(3), 407–421. https://doi.org/10.1111/j.2044-8295.1997.tb02648.x
Yum, Y. N., & Law, S. (2019). Interactions of age of acquisition and lexical frequency effects with phonological regularity: An ERP study. *Psychophysiology*, *56*(10). https://doi.org/10.1111/psyp.13433
Zeno, S. M., Ivens, S. H., Hillard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide.* Brewster, NJ: Touchstone.
Zevin, J. D., & Seidenberg, M. S. (2002). Age of Acquisition Effects in Word Reading and Other Tasks. *Journal of Memory and Language*, *47*(1), 1–29. https://doi.org/10.1006/jmla.2001.2834
Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition*, *32*(1), 31–38. https://doi.org/10.3758/BF03195818