# Age of acquisition effects on the decomposition of compound words

Elsherif, M.M.; Catling, J.C.

*DOI:*
10.1080/20445911.2021.2013246

# Title: Age of Acquisition Effects on the decomposition of compound words

Elsherif, M.M. *[1], 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

**Abstract**

Age of acquisition (AoA) is a measure of learning experience and a strong predictor of lexical retrieval. According to the integrated view, the AoA effect results from the development of semantic representations and the mappings between these representations. This has not been considered in morphologically complex words. The integrated account predicts that the AoA effect should be larger in tasks requiring greater semantic processing and any AoA effects should be shown in the early processes of word recognition. The present study investigates these predictions in compound words, which differ from monomorphemic words in terms of ease of mapping and semantic processes in lexical retrieval. Forty-eight participants completed a compound lexeme segmentation (CLS) task, in which participants named either the head or modifier depending on the number above the compound word, to establish how semantics are involved in processing the head and the modifier. The results demonstrated that semantics influenced the naming of the modifier to a greater extent than the head, with the AoA effect being larger in the modifier than the head. Our findings provide evidence that aligns with the multiple origins of AoA effects in the language processing system.

Keywords: age of acquisition; compound lexeme segmentation; compound word; lexical retrieval; morphology; semantics

Linguistic experience is known to modulate the speed of acquiring and processing language, which, in turn, affects cognition (see review by Perfetti, 2007). One measure of linguistic experience is the age at which a word or an object is first encountered and learned such that early acquired words are processed more quickly and accurately than later acquired words (i.e. the Age of Acquisition effect: AoA). This effect has been demonstrated in typical and atypical populations, computational modelling, neuroimaging methodologies and a variety of behavioural tasks including object naming, face naming, visual duration threshold, text reading, word naming and lexical decision across different languages (e.g. Barry et al., 1997; Catling et al., 2013; Catling & Johnston, 2006a, 2006b, 2006c; Dent et al., 2007; Gerhand & Barry, 1999; Juhasz & Rayner, 2003, 2006; Lewis, 1999; Moore & Valentine, 1998; Morrison & Ellis, 1995; Sohrabi, 2018; Yum & Law, 2019 see also reviews by Brysbaert & Ellis, 2016; Johnston & Barry, 2006; Juhasz, 2005).

***Theories of the AoA effect***

There are two predominant theories that explain the processes that delineate the AoA effect. The semantic theory, originally presented by van Loon-Vervoon (1989, cited in Brysbaert et al., 2000), posits that early-acquired words are placed at the hub (or centre) of the network, thus developing more semantic connections to other concepts, affording early-acquired words richer semantic representations than their later-acquired counterparts. This enables early-acquired words to be easily processed, more accessible and less likely to be forgotten (see reviews by (Brysbaert & Ellis, 2016; Henry & Kuperman, 2013; Marful et al., 2016; Steyvers & Tenenbaum, 2005). Put simply, the AoA effect results from gradual development of semantic representation. Evidence has demonstrated that the more semantic processing involved in the task, the stronger the AoA effect (e.g. AoA effects are larger for LDT than for word naming; Catling & Johnston, 2009), as word naming requires access from orthography to phonology without recourse to semantics, while LDT requires access from orthography to semantics (Cortese et al., 2018; Kuperman, 2013).

An alternative, but not mutually exclusive theory is the arbitrary mapping (AM) hypothesis (e.g. (Ellis & Lambon Ralph, 2000; Lambon Ralph & Ehsan, 2006; P. Monaghan & Ellis, 2010). According to this hypothesis, early-acquired words benefit from the rich resources available in the neural network, leading to these words being consolidated in the mental lexicon. As a result, early-acquired items modify the connections between input and output representations, causing the network to lose its rich resources for consolidation of the lexical representation. Late-acquired items thus need to be fitted into the network's structure, which is easy when the relationship between orthography/phonology and semantics is systematic and regular (e.g. word naming) but when it is arbitrary (e.g. picture naming), late-acquired words incur a processing cost (Zevin & Seidenberg, 2002). Evidence has demonstrated that AoA effects are stronger for tasks that assess the arbitrary relationship between orthography-to-phonology mappings than systematic and regular orthographic-to-phonological mappings (e.g. Lambon Ralph & Ehsan, 2006). This is because later acquired words cannot 'gain' from any general rules that have built up due to regularity of input – output coding (Lambon Ralph & Ehsan, 2006).

The aforementioned theories have been merged to explain the AoA effect ((Brysbaert & Ellis, 2016; Catling et al., 2021; Catling & Elsherif, 2020; Chang et al., 2019; Chang & Lee, 2020; Cortese et al., 2020; Dirix & Duyck, 2017; Menenti & Burani, 2007). According to this integrated account, the AoA effect results from a combination of the formation of representations and the changing plasticity in the neural network throughout development (Chang et al., 2019). Chang and Lee (2020) observed that the AoA effect was stronger in character naming than the lexical decision task. The former was more affected by semantics than the latter, while the former showed an interaction of regularity and AoA such that the AoA effect was stronger in inconsistent than in consistent words. This supports the integrated account of the AoA effect, as early-acquired words benefit from more connections and greater accessibility than late-acquired words, such that lexical processing is shaped gradually by the experience of learning during development (Brysbaert & Ellis, 2016; Dirix & Duyck, 2017). This theory has been limited to monomorphemic items and can be strengthened by generalising the pattern of findings to morphologically complex words such as compound words. By assessing compound words, one can examine the involvement of lexical and semantic elements during lexical and lexemic processing.

***AoA in compound words***

Compound words (e.g. airplane) are formed of two lexemes. A compound word obtains its semantic category and morphosyntactic features of the whole compound from the head lexeme (i.e. *airplane* is a plane), which also makes the whole compound a noun. Different languages vary as a function of whether their compounds are left- or right-headed. With a few exceptions, English compounds are right-headed (i.e. the second or rightmost constituent determines the semantic category and morphosyntactic features of the whole compound) (Plag, 2018; Williams, 1981). The modifier on the other hand specifies the meaning of the head noun (e.g. a *airplane* is a specific type of plane dedicated to air travel; Benczes, 2005, 2014, 2015; Günther et al., 2020). This function of adding specificity narrows the semantic domains. The influence of semantics on word naming in monomorphemic words is unclear (see Balota et al., 2004; Kuperman, 2013; J. Monaghan & Ellis, 2002) but in comparison to compound words, semantics is argued to play an important role. According to Cortese and Schock (2013), disyllabic words such as compound words have irregular spelling-to-sound correspondence, thus readers need more time to compute the pronunciation of the word, leading to larger semantic effects than monomorphemic words, which have regular spelling-to-sound correspondence.

Very few studies have assessed the role of the AoA effect in compound words. Juhasz et al. (2015) observed that rated AoA and word frequency of the compound and only frequency, not AoA, of the lexemes affected word naming latencies (see also Juhasz, 2018 who replicated the same pattern of findings for the AoA effect in eye-tracking). Juhasz et al. concluded that the AoA effect may occur at the lexical/semantic level and that the semantic representations of the lexemes are not automatically activated when the compound word is processed (see also Kuperman, 2013) but this does not mean that lexical decomposition does not occur, as lexeme frequency affects word naming (e.g. Juhasz et al., 2003, 2015). Elsherif et al. (2020) observed that the AoA of the compound word contributed to naming latencies in both word naming and combinatorial naming (i.e. reading two lexemes as one word), while the AoA of the lexemes only contributed to the naming latencies of combinatorial naming. The authors concluded that the AoA effect originates at the lexical level but does not necessitate access to semantics. However, these findings have not considered the influence of the AoA effect on the head and modifier lexemes.

The present study was designed to investigate the integrated view of the AoA effect using compound lexeme segmentation (CLS). Before we delve into the CLS, this task is similar to the phoneme elision task, in which increasingly smaller segments are removed from the stimulus at increasingly higher levels of linguistic complexity, from words down to phonemes within clusters (Wagner et al., 1999). In this task, participants verbally produce a novel word after a particular phoneme has been deleted (e.g., /kəp/ without the /k/ sound produces /əp/, or UP). Phoneme elision has been shown to be a robust measure of phonological awareness. In addition, phoneme elision has been shown not to correlate with semantic variables such as vocabulary composite (e.g. Elsherif et al., 2021a, 2021b, 2021c). In a similar fashion, the CLS task requires the participant to name only the head or modifier of the compound word dependent on the number above the compound word. For instance, if the number one was placed above the word "airplane", "air" would be named. This requires the elimination of a lexeme from a compound word such as "plane" from "airplane", if a number one is positioned above the latter in order to produce "air". This needs the person to accurately encode the compound word in their working memory. To produce the revised lexeme, the person must be able to verbally rehearse the compound word accurately, and hold it in their memory long enough to allow for accurate segmentation and manipulation of it, all within a timeframe that allows for accurate production of the revised lexeme.

The logic behind the CLS task is that it makes the reader extract the individual lexeme, thus forcing the reader to travel via the decomposition route, whilst inhibiting the direct lexical route. Brooks and de Garcia (2015) assessed the processing of transparent compounds (e.g. roadside), opaque compounds (e.g. butterfly), and morphologically simple words (e.g. spinach). They used a word naming task, which involved partial-repetition priming and recorded the neural activity of participants. This analysis was limited to visual processing of compound word primes. Their results showed that, compared to the processing of morphologically simple words, processing of transparent compounds was related to greater neural activation in the anterior middle temporal gyrus (around 250-470ms), with stronger effects occurring in the posterior superior temporal gyrus (430-600ms). The authors concluded that compound processing involved a decomposition stage independent of semantics. We suggest that if these findings are brought across to a CLS task, that an AoA effect of the compound word should be observed when naming both the modifier and head lexeme and that there should be no interaction between AoA and the individual lexemes, thus supporting the mapping theory of the AoA effect.

Secondly, CLS enables us to assess the influence of the AoA effect on lexemic processing. This task allows us to extract the modifier or head, which are processed in different ways. It is unclear whether the meaning of the compound word is located in the head or the modifier lexeme, Kuperman (2013) investigated the non-relational semantic properties of compound word recognition and whether the whole compound word or the meanings of the lexemes are accessed for compound word recognition. They used rated imageability, valence, arousal and concreteness of the compound word and its respective lexemes to assess which of the measures would drive the lexical decision times obtained from the English Lexicon Project (Balota et al., 2007). The authors observed that when adding a baseline model containing word length, individual lexeme frequency and compound word frequency, the valence of both lexemes affected compound processing. In combinatorial naming (i.e. name two lexemes as a compound word), Elsherif et al. (2020) observed that the rated imageability (a semantic variable) of the head and modifier contributed to naming latencies. As a result, it could be concluded that there may be equal effects of both lexemes for meaning composition of the compound word. If this is the case, the AoA effects of the compound word should be similar for the head and the modifier lexemes.

However, evidence has repeatedly demonstrated that the involvement of semantics in compound words is driven by the modifier lexeme (e.g. Günther et al., 2020; Günther & Marelli, 2019)[1]. Gunther and Marelli (2019) investigated semantic effects on simple lexical decision times in a large set of English compounds (Balota et al., 2007) and observed that the modifier composition (i.e. similarity between the modifier of the lexeme and the meaning of the compound word, together with the ease of integrating the modifier meaning into a combined concept; Gagné & Spalding, 2009; Günther & Marelli, 2016) was a stronger predictor than the head composition (i.e. analogous to the modifier composition) in predicting lexical decision times. In Experiment 1, Gunther et al. (2020) used a lexical decision task and changed one grapheme in the first, middle or final position in one of the constituents for each item of the compound word (e.g. <airport> became <airpurt>). Gunther et al. noted that the head composition and modifier composition contributed to rejection times. However, similar to Gunther and Marelli (2016, 2019), the modifier composition was a stronger predictor than the head composition in predicting the rejection times in lexical decision tasks. In Experiment 2, Gunther et al. used pseudo-compounds instead of modifying graphemes in compound words (e.g. Ritterglocke means knightbell) and noted an effect of the modifier composition in a lexical decision task, whereas participants took longer to reject novel compounds when the modifier was more easily integrated into compound meanings. These findings indicate that processing speed is affected by the ease of integrating the modifier of the compound word into a combined concept, thus the modifier contributes more to the semantics of the compound word than the head lexeme (see also supporting evidence in corpus linguistics, in which evidence has shown that the role of the modifier contributes to the meaning of the compound word more strongly than the head lexeme; Elsherif & Winter, 2021). It is therefore expected that the modifier lexeme affects the meaning composition of the compound word. If this is the case and following the representation theory, we predict that the

[1] Like English, compound words in German are right-headed.

AoA of the modifier should be shown only when naming the modifier if the AoA effect is driven solely by semantics. There should also be an interaction of the rated AoA of the compound word and position such that the magnitude of the AoA effect would be larger in naming the modifier than the head lexeme. The mapping hypothesis would not predict this interaction. The integrated account predicts that the rated AoA of the compound word will be observed in naming the head and modifier lexemes and there should be an interaction between the rated AoA and position of the compound word.

## *Method*

### Participants

To reduce experimenter bias, the data were analysed after all the participants were recruited and a stopping rule was introduced. Based on Elsherif et al.'s (2020) experiment, we used 48 British monolingual undergraduate students aged 18-20 years (M = 18.16±0.35 years; 3 males), who were given course credits as remuneration. A post-hoc power analysis, using Westfall et al.'s (2014) formula for effect size calculation (see Brysbaert and Stevens, 2018, for discussion), indicated that our sample size exceeded the number required to reach the desired level of power of 0.80 (minimum of 13 recommended, while we included the data from 48 participants in the analyses)[2]. The experiment was conducted in accordance with the British Psychological Society's ethical guidelines and approved by the University's ethical committee. All participants had normal or corrected-to-normal vision and signed a consent form to participate in the study.

### Materials

We used the same stimuli as Elsherif et al. (2020), which we briefly describe in this section. Each participant was shown 226 words that were primarily noun-noun compounds (see appendix). Word frequencies as Zipf values were extracted from the SUBTLEX-UK database for the compound word

[2] Westfall et al. (2014) published a theoretical analysis of a mixed effects model and created a website ((https://jakewestfall.shinyapps.io/two_factor_power/), which allows researchers to calculate the power of an experiment and the number of items/participants required for a well-powered experiment. We estimated the sample size using the following values: Effect size *d*: 0.063, residual variance: 0.61, participant intercept variance: 0.23, item intercept variance: 0.16, participant-by-item variance: 0.219, participant slope variance: 0, item slope variance: 0.

and modifier and head lexemes of the compound word (van Heuven et al., 2014). Letter length, AoA, semantic transparency (ST: i.e. the meaning of the compound word is related either to the modifier or the head), and lexeme meaning dominance (LMD: i.e. the degree to which the meaning of a compound word is contained in either the head or the modifier) were taken from Juhasz et al.'s (2015) database. Familiarity measures for each of the lexemes, all based on subjective ratings, were taken from their respective databases (familiarity: Balota et al., 2001). The AoA of each lexeme was collected from Cortese and Khanna (2008) and Schock et al. (2012) (see Table 1 for psycholinguistic characteristics). The mean length of our stimuli was divided into modifier length and head length. As participants had to name either the first or second lexeme, it was more appropriate to include the first and second lexeme length only.

[Insert Table 1]

**Procedure**

Participants were tested individually in front of a computer screen with a microphone approximately 15cm away from the mouth. Using E-prime software to collect the responses (E-studio, E-Prime 2.0), participants were instructed to say either the first or second lexeme of the compound word, depending on the number above the compound word, as fast as possible without compromising their accuracy. For example, if a number one was placed on top of the compound word "airplane", the participant must say "air" and if a number two was placed above, the participant must say "plane". A fixation cross appeared at the centre of the screen for 250ms, after which the stimulus appeared in the same position, which was shown until the lexeme was named or 2000ms had passed. Stimuli were presented in upper case using Arial font (size: 34). This was followed by an inter-trial interval of 1000ms. Each session lasted approximately ten minutes.

**Analysis**

All analyses were conducted in R 3.6.1. (R Core Team, 2019), using the "tidyverse" package version 1.3.0 for data processing (Wickham et al., 2019). We analysed the data using linear mixed-effect models (LMM) using the lme4 package, version 1.1.26 (Bates et al., 2015). T values were computed for each variable of interest and a variable was significant at the alpha = .05 level if the absolute *t* value was greater than 1.96 (Baayen et al., 2008). In model summaries, we report both marginal ($R^2m$) and conditional $R^2$ values ($R^2c$), using the 'MuMIn' package version 1.43.17 (Barton & Barton, 2015). The former is an estimate of the variance explained by fixed factors only, whereas the latter explains the variance of the whole model (i.e. fixed and random factors; Nakagawa & Schielzeth, 2013). Prior to analysis, error rates and missed/late responses were removed. Outliers including responses faster than 200ms and greater than or below 2SD from the group median were also removed, leading to the removal of 4% of the responses.

The following LMM analyses were conducted separately on the modifier and head in CLS, with item and subject as random factors and all the predictors as fixed factors. All the predictors were centred on their means and the RTs were inverse-transformed (Brysbaert & Stevens, 2018). We used a benchmark model that included compound word, the frequency of the modifier and head, and the initial phonetic onset (Elsherif et al., 2020; Juhasz et al., 2015; Kuperman, 2013). In contrast to Elsherif et al., Juhasz et al. and Kuperman, the baseline model removed the whole word length and only included the length of the modifier and head, as participants only had to name the respective lexeme. In addition, to control for potential onset differences, acoustic properties of onset were also included and coded dichotomously (1 or 0; Balota et al., 2004; Spieler & Balota, 1997; Treiman et al., 1995). For all LMM models, collinearity diagnostic analyses showed a variance inflation factor (VIF) of 6.15, which was driven by the alveolar initial phoneme onset. Removing the alveolar initial phoneme onset from the analyses[3], produced a VIF of around 1.55. In addition, we checked the extent to which the AoA of the compound word and its lexemes could explain the variance beyond the main linguistic processing predictors (without the fear of collinearity). Each

[3] The patterns of the results did not change when alveolar was included or excluded as a fixed predictor.

variable of interest was added separately to the model. The correlation matrix between the variables is shown in Table 2.

After the analysis was conducted for each position, we examined the influence of lexical-semantics or regularity and AoA on both positions by conducting a LMM analysis, combining position data into a dependent variable. The combined model included the variable position (-1 for second and 1 for first position) as an additional predictor, together with the fixed and random factors. Following this, three interactions were assessed using a conservative nested model comparison approach (Barr et al., 2013). The significance was assessed by determining whether the model fit improved significantly by applying a likelihood ratio test comparison between models with and without the interaction of interest.

[Insert Table 2 about here]

***Results***

Following the advice of Manly and Wells (2015) and Woods et al. (2021), across the samples the rate of missingness ranged around 2% for gender but it was missing at random. The data and materials for all experiments are available at https://osf.io/g5b3m/. In the baseline model for the CLS, two out of nine initial phoneme onset features (i.e. bilabial and labiovelar) were significant predictors of the naming latencies of the modifier lexeme, whereas five out of nine initial phoneme onset features (i.e. bilabial, labiodental, postalveolar, glottal and velar) significantly contributed to the naming latencies of the head lexeme, thus demonstrating the onset effect in CLS (see Tables 3 and 4). The familiarity, AoA and LMD of the compound word, and the AoA of the modifier lexeme made significant contributions to the naming latencies of the modifier lexeme. The frequency of the compound words and lexemes, together with the ST of the compound word and length and familiarity of the lexemes, were not significant predictors of the naming latencies for the modifier of the CLS. For the head, the AoA, ST and LMD of the compound word were significant, unlike the familiarity, word frequency and word length of the compound word and lexemes. The estimate value for familiarity and AoA, which has been argued to be a measure of lexical/semantic processes in compound word processing (e.g. Juhasz et al., 2015), was also numerically larger for the modifier lexeme than the head lexeme (Tables 5 and 6), while the LMD predictor was similar between positions, indicating that lexical-semantic effects are stronger for the modifier than for the head lexeme.

[Insert Tables 3-6 about here]

**Interaction of the AoA of the compound word and position**

The combined model was associated with $R^2m$ = 0.53% and $R^2c$ = 24.7% where position was a significant predictor (*b* = 0.03, SE = 0.004, *t* = 8.30), with faster RTs for the modifier (M = 750ms, SD = 179) than for the head (M = 770ms, SD = 183)[4]. Position by frequency, familiarity or AoA of the compound word was added into the combined model separately as a fixed factor. Adding position x frequency to the model resulted in a significant improvement ($\chi 2(1) = 4.31$, $p = .04$), where $R^2m$ = 0.66% and $R^2c$ = 25.4% respectively. Adding position x familiarity to the model also resulted in a significant improvement ($\chi 2(2) = 7.81$, $p = .02$), where $R^2m$ = 0.71% and $R^2c$ = 25.4% respectively. A similar result was obtained for AoA and position ($\chi 2(2) = 16.77$, $p < .001$), where $R^2m$ = 0.80% and $R^2c$ = 25.4% respectively. The interaction patterns (Figure 1) showed that all target effects were stronger for the modifier than the head, despite the fact that the RTs were longer for the head than the modifier. Following these interactions, the models were split into two sub-models: the modifier lexeme and the head lexeme. Position was removed from the equation and the same procedures for the analyses from the reduced model were applied to the sub-models. In the head lexeme sub-model, the predictor of AoA for the compound word was the only significant predictor, while in the modifier lexeme sub-model, the predictor of AoA and familiarity for the compound word were

[4] Although the head was more frequent, early-acquired and shorter in letter length than the modifier of the compound word, suggesting that the former should be processed more quickly than the latter, the opposite was observed.

significant predictors of naming the modifier. The predictor of word frequency for the compound word were not evident in naming the modifier or head lexeme[5].

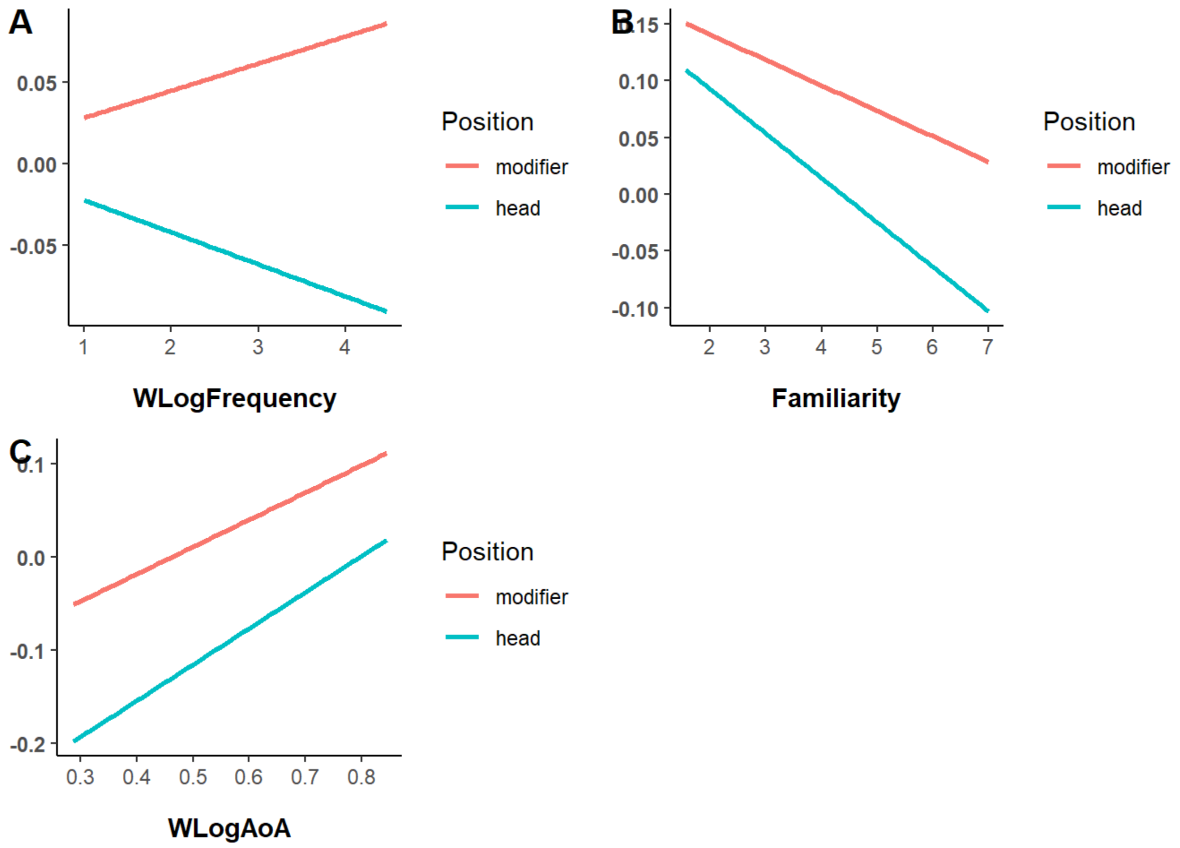


**Figure 1**. The interaction patterns between (A) position and frequency, (B) position and familiarity and (C) position and AoA.

## Discussion

This study investigated the AoA effects using CLS. We found that the AoA and familiarity of the compound word had the expected effect for CLS in both positions. The study showed that compounds that were early-acquired and familiar were named faster than those acquired later and unfamiliar, independent of other lexical variables. In addition, we predicted that the AoA of the

[5] The different patterns that emerge for word frequency of and familiarity of the compound word may occur, as they do overlap in terms of being frequency measures. However, the database used in this study could drive the effect. The word frequency measure was extracted from SUBTLEX-UK database, while we extracted the familiarity of the compound word from Juhasz et al.'s (2015) database. Juhasz et al. asked participants to rate the familiarity of the compound word based on its meaning and frequency, thus the familiarity measure may not only be affected by subjective frequency but also meaning, while the word frequency is a measure of objective frequency. In addition, the correlation between these two measures was weak to moderate ($r$ = . 33), thus supporting the argument that they may share a frequency-component but perhaps differ in terms of semantics, leading to the different pattern of findings.

modifier, together with the compound word, should be observed when naming the modifier, as the modifier is more affected by semantics, which is in line with previous results (Elsherif et al., 2020). We observed that AoA of the compound word contributed to the naming latencies of the modifier lexeme and the AoA of the modifier lexeme contributed to the naming of the modifier. The AoA of the compound word, not that of the head lexeme, contributed to the naming of the head lexeme of the compound word. In addition, we investigated the interaction between the effect of lexical-semantic variables such as familiarity and position in the CLS, observing that the lexical-semantic effects were stronger for the modifier than the head lexeme. The magnitude of the AoA effect was also found to be stronger in the modifier than the head lexeme. This result suggests that the AoA effect has a common origin with lexical-semantic variables, as predicted by the semantic theory (see review by Brysbaert & Ellis, 2015).

The effect of lexeme AoA was observed in the modifier of the compound word, while the lack of the AoA in the head of the compound word partially supports Elsherif et al. (2020), who argued that the presence of the lexeme AoA occurs when semantics is involved. The AoA of the modifier indicates that semantics is involved in naming the modifier, while the lack of the AoA effect in the naming of the head lexeme suggests that semantics is not involved in naming the head lexeme. In addition, the AoA effect of the compound word decreases from the modifier to the head. This indicates that the AoA effect for the head occurs in the mappings between representations, with semantics contributing a minimal role. These findings demonstrate that the AoA effect cannot be determined solely by the mappings between representations but also by the formation of the semantic representation. Also, the AoA effect influences language processing in terms of both the connections between levels of representation and the actual levels of representation themselves. Overall, the findings of AoA effects in compound words follow the prediction of the integrated view of AoA (Brysbaert & Ellis; Chang et al., 2019; Dirix & Duyck, 2017; Menenti & Burani, 2007).

Different theories have been proffered to explain the AoA effect. The AM hypothesis argues that the AoA effect results from a decrease in the resources of the neural network as more words are acquired, thus late-acquired words that have arbitrary links between representations should face a larger cost, leading to the AoA effect (Ellis & Lambon Ralph, 2000). Alternatively, AoA effects can be related to semantic representations, such that early-acquired words have richer representations and more connections with other concepts (Henry & Kuperman, 2013; Steyvers & Tenebaum, 2005). However, these hypotheses are not mutually exclusive and the levels both within and between representations could influence lexical-semantic processing, producing AoA effects (see review by Brysbaert & Ellis, 2016). We observed that the familiarity and AoA effects of the compound word were stronger in naming the modifier than the head lexeme. The AoA of the modifier (not the head) lexeme also contributed to the naming latencies of a specific lexeme, indicating that the AoA effect has a semantic origin. The AoA of the compound word was also shown in the head lexeme, in which semantics plays a minimal role, indicating that the AoA effect is also likely to occur in the mapping between representations. This indicates that the AoA effect influences language processing in terms of both the connections between levels of representation and the actual levels of representation themselves.

The CLS approach used in this study is theoretically important, not only for AoA effects but also for investigating lexemic decomposition. The results of the CLS established that the influence of lexical-semantic variables differs between the head and modifier in that these effects were greater in the modifier than the head. The directionality of these effects is shown, together with the AoA and familiarity of the modifier, indicating that lexical-semantic and lexemic processing occurs early in the processing of the compound word. It is likely that this finding occurs for two reasons: firstly, compound words are naturally long words, thus the relationship between spelling and sound are more arbitrary, causing semantics to be involved (Cortese et al., 2013) and secondly, the modifier changes the relational meaning of the head for the compound word (e.g. a plane made for air, a plane made for sea etc.; Benzes et al., 2005, 2014, 2015), which is more likely to be arbitrary, and tries to constrain the meaning of the head of the compound word. It is important to state that the effect of the CLS for the interaction of the AoA effect and position was small but reliable. Larger differences could be observed when using tasks that are arbitrary in nature such as a CLS version of picture naming, which should lead to larger AoA effects. It is important to "note that as the sample [size]…was moderate" (Elsherif et al., 2017, p. 26), we should still remain cautious about these effects. A large-scale and systematic comparison of the CLS with picture naming and word naming is a key topic for future investigation to assess further the integrated view of the AoA effect, perhaps using a using 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).

In terms of reading compound words, the CLS is an elision task. Elision tasks are usually found to be more strongly correlated with reading fluency than other phonological tasks (e.g. blending), which is in line with earlier findings of Wagner et al. (1999). In the current study, the high performance found in the CLS task may be indicative of a development of skilled reading associated with a deeper awareness of morphological units, which in turn, allows for more efficient segmentation, deletion and blending of lexemes and phonemes, complementing Stanovich (1992). The high accuracy in CLS may be indicative of better developed morphological processing and conceptualisation skills. In addition, the high accuracy may be indicative that it is difficult to suppress and inhibit the direct lexical route during reading, as the processing of compound words is perhaps lexical in nature. Future research should use this task to assess the development from morphological route to the lexical route to assess how the morphological processes may develop in reading.

To conclude, this study investigated the AoA effects by conducting a CLS task. Although compound words are processed differently to monomorphemic words, we found that the AoA of the compound word was larger for naming the modifier than the head lexeme. In addition, the AoA of the compound word was found to be present in both the modifier and head lexeme and the AoA of the modifier, not the head, affected their respective lexeme. This suggests that the AoA effect contributes to the gradual development of lexical representation and the relationship between representations.

## Data availability statement

The data and materials for all experiments are available at https://osf.io/g5b3m/.

## References


Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

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., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*(4), 639-647. https://doi.org/10.3758/BF03200465

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

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

Barton, K., & Barton, M. K. (2015). *Package 'mumin'*. *Version*, *1*, 18. https://cran.r- project.org/web/ packages/MuMIn/MuMIn.pdf.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Benczes, R. (2005). Metaphor-and metonymy-based compounds in English: a cognitive linguistic approach. *Acta Linguistica Hungarica*, *52*(2-3), 173-198. https://doi.org/10.1556/aling.52.2005.2-3.3

Benczes, R. (2014). Repetitions which are not repetitions: the non-redundant nature of tautological compounds. *English Language and Linguistics*, *18*(3), 431-47. https://doi.org/10.1017/S1360674314000112

Benczes, R. (2015). Are exocentric compounds really exocentric?. *SKASE Journal of Theoretical Linguistics*, *12*(3), 54-73.

Brooks, T. L., & Cid de Garcia, D. (2015). Evidence for morphological composition in compound words using MEG. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00215

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., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), 1–20. https://doi.org/10.5334/joc.10

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

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

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*, *29*(5), 662–674. https://doi.org/10.1080/09658211.2021.1931695

Catling, J. C., South, F., & Dent, K. (2013). The Effect of Age of Acquisition on Older Individuals with and without Cognitive Impairments. *Quarterly Journal of Experimental Psychology*, *66*(10), 1963–1973. https://doi.org/10.1080/17470218.2013.771689

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

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

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *Quarterly Journal of Experimental Psychology*, *66*(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

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

Dirix, N., & Duyck, W. (2017). 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

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

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., Saban, M. I., & Rotshtein, P. (2017). The perceptual saliency of fearful eyes and smiles: A signal detection study. *PLOS ONE*, *12*(3), e0173199. https://doi.org/10.1371/journal.pone.0173199

Elsherif, M. M., Wheeldon, L., & Frisson, S. (2021a). *Orthographic precision for word naming in skilled readers* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/gpk9m

Elsherif, M. M., Wheeldon, L. R., & Frisson, S. (2021b). Do dyslexia and stuttering share a processing deficit? *Journal of Fluency Disorders*, *67*, 105827. https://doi.org/10.1016/j.jfludis.2020.105827

Elsherif, M. M., Wheeldon, L. R., & Frisson, S. (2021c). Phonological precision for word recognition in skilled readers. *Quarterly Journal of Experimental Psychology*, 17470218211046350. https://doi.org/10.1177/17470218211046350

Elsherif, M. M., & Winter, B. (2021). *What are English compounds made of?* [Preprint]. PsyArXiv. https://psyarxiv.com/xpnds/

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language*, *60*(1), 20–35. https://doi.org/10.1016/j.jml.2008.07.003

Gerhand, S., & Barry, C. (1999). 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

Günther, F., & Marelli, M. (2016). Understanding Karma Police: The Perceived Plausibility of Noun Compounds as Predicted by Distributional Models of Semantic Representation. *PloS One*, *11*(10), e0163200. https://doi.org/10.1371/journal.pone.0163200

Günther, F., & Marelli, M. (2019). Enter sandman: Compound processing and semantic transparency in a compositional perspective. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *45*(10), 1872–1882. https://doi.org/10.1037/xlm0000677

Günther, F., Marelli, M., & Bölte, J. (2020). Semantic transparency effects in German compounds: A large dataset and multiple-task investigation. *Behavior Research Methods*, *52*(3), 1208–1224. https://doi.org/10.3758/s13428-019-01311-4

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

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

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., 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., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British journal of psychology*, *94*(2), 223-244. https://doi.org/10.1348/000712603321661903

Kuperman, V. (2013). Accentuate the Positive: Semantic Access in English Compounds. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00203

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

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

Manly, C. A., & Wells, R. S. (2015). Reporting the Use of Multiple Imputation for Missing Data in Higher Education Research. *Research in Higher Education*, *56*(4), 397–409. https://doi.org/10.1007/s11162-014-9344-9

Marful, A., Gómez-Ariza, C. J., Barbón, A., & Bajo, T. (2016). Forgetting “Novel” but Not “Dragon”: The Role of Age of Acquisition on Intentional and Incidental Forgetting. *PLOS ONE*, *11*(5), e0155110. https://doi.org/10.1371/journal.pone.0155110

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

Monaghan, J., & Ellis, A. W. (2002). 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

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

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

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Plag, I. (2018). *Word-formation in English*. Cambridge University Press.

Schock, J., Cortese, M. J., Khanna, M. M., & Toppi, S. (2012). Age of acquisition estimates for 3,000 disyllabic words. *Behavior Research Methods*, *44*(4), 971-977. https://doi.org/10.3758/s13428-012-0209-x

Sohrabi, A. (2018). Age of Acquisition Effect: Evidence from Single Word Reading and Neural Networks. *Basic and Clinical Neuroscience Journal*. https://doi.org/10.32598/bcn.9.10.120

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, *8*(6), 411-416. https://doi.org/10.1111/j.1467-9280.1997.tb00453.x

Stanovich K.E. (1992). Speculations on the Causes and Consequences of Individual Differences in Early Reading Acquisition. In: P.B. Gough, L.C. Ehri, & R. Treiman (eds). *Reading Acquisition*. Lawrence Erlbaum Associates, Inc., 307–342.

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

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*, 104060. 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. http://dx.doi.org/10.1037/0096-3445.124.2.107

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

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (1999). *Comprehensive test of phonological processing: CTOPP*. Austin, TX: Pro-ed.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Williams, E. (1981). Argument Structure and Morphology. *The Linguistic Review*, *1*(1). https://doi.org/10.1515/tlir.1981.1.1.81

Woods, A. D., Davis-Kean, P., Halvorson, M. A., King, K. M., Logan, J. A. R., Xu, M., … Elsherif, M. M. (2021, November 3). Best Practices for Addressing Missing Data through Multiple Imputation. https://doi.org/10.31234/osf.io/uaezh

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

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

| **Predictors** | Compound word | | | 1st lexeme | | | 2nd lexeme | | |
|---|---|---|---|---|---|---|---|---|---|
| | M | SD | Range | M | SD | Range | M | SD | Range |
| Word length | 8.61 | 1.34 | 6-13 | 4.42 | 1.02 | 2-8 | 4.19 | 0.84 | 2-7 |
| Frequency (out of 7) | 2.67 | 0.72 | 0.696-4.48 | 4.50 | 0.79 | 2.69-6.78 | 4.69 | 0.82 | 2.38-7.42[a] |
| Familiarity (out of 7) | 5.77 | 1.15 | 1.57-7.00 | 4.72 | 1.66 | 1-6.41 | 4.98 | 1.50 | 1-6.43 |
| AoA (out of 7) | 4.70 | 1.22 | 1.93-7.00 | 3.54 | 0.86 | 1.70-6.10 | 3.35 | 0.82 | 2-6.30 |
| ST (out of 7) | 4.59 | 1.33 | 1.6-6.71 | NA | | | NA | | |
| LMD (out of 10) | 5.17 | 1.42 | 1.47-8.67 | NA | | | NA | | |

**Table 1.** Descriptive statistics for word target characteristics for compound word.

**Note:** AoA = Age-of-Acquisition, ST = Semantic Transparency and LMD = Lexeme Meaning Dominance.
[a] Although there is a discrepancy between the maximum and range of values shown in this table, this discrepancy is from the van Heuven et al.'s SUBTLEX-UK online database and the Likert scale used (scale of 1-7 discussed on their website). We had used the word "to", which had a Zipf scale of 7.42 and function words tend to go beyond the maximum score.

**Table 2.** Correlation between independent variables. CF = compound word frequency, CFA = Compound word familiarity, CAoA = Compound age-of-acquisition, ST = semantic transparency, LMD = lexeme meaning dominance, ML = The modifier lexeme of length, MLF = The modifier lexeme of frequency, MAoA = The modifier lexeme of age-of-acquisition, MFA = The modifier lexeme of familiarity, HL = The head lexeme of length, HF = The head lexeme of frequency, HFA = The head lexeme of familiarity, HAoA = The head lexeme of age-of-acquisition. $^{+}p < .10$, * $p < .05$, ** $p < .01$ and *** $p < .001$.

| | CF | CFA | CAOA | ST | LMD | ML | MF | MFA | MAoA | HL | HF | HFA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CFA | 0.33*** | | | | | | | | | | | |
| CAoA | -0.20** | -0.65*** | | | | | | | | | | |
| ST | -0.08 | 0.37*** | -0.37*** | | | | | | | | | |
| LMD | -0.01 | 0.00 | -0.06 | 0.03 | | | | | | | | |
| ML | -0.02 | 0.05 | -0.02 | 0.04 | -0.07 | | | | | | | |
| MF | 0.27*** | 0.22*** | -0.11 | 0.11 | -0.10 | -0.11 | | | | | | |
| MFA | 0.09 | 0.19** | -0.17** | 0.21** | -0.09 | 0.08 | 0.49*** | | | | | |
| MAoA | -0.09 | -0.15* | 0.30*** | -0.16* | 0.00 | 0.28*** | -0.59*** | -0.50*** | | | | |
| HL | 0.04 | 0.14* | 0.09 | 0.08 | 0.01 | 0.03 | 0.00 | 0.00 | 0.03 | | | |
| HF | 0.15* | 0.00 | -0.01 | 0.09 | -0.17** | 0.02 | 0.21*** | 0.13+ | -0.06 | -0.21*** | | |
| HFA | 0.09 | 0.10 | -0.02 | 0.23*** | 0.06 | 0.03 | 0.05 | 0.09 | -0.02 | -0.02 | 0.55*** | |
| HAoA | 0.01 | -0.11 | 0.08 | -0.08 | -0.14* | 0.02 | -0.05 | -0.05 | 0.11 | 0.24*** | -0.55*** | -0.41*** |

**Table 3.** Linear mixed effects regression results for the baseline model for the modifier lexeme[a].

| Values | ML | CFreq | MFreq | HFreq | Bilabial | Labiodental | Dental | Labiovelar | postalveolar | PA | Glottal | Velar | Voiced |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *β* | -0.001 | -0.007 | 0.003 | -0.005 | 0.020 | 0.016 | 0.003 | 0.019 | 0.018 | -0.016 | 0.011 | 0.005 | 0.009 |
| SE | 0.005 | 0.005 | 0.005 | 0.005 | 0.007 | 0.010 | 0.026 | 0.008 | 0.016 | 0.014 | 0.009 | 0.008 | 0.006 |
| 2.5% CI | -0.010 | -0.016 | -0.007 | -0.015 | 0.006 | -0.004 | -0.048 | 0.003 | 0.014 | -0.044 | -0.006 | -0.010 | 0.023 |
| 97.5% CI | 0.009 | 0.003 | 0.013 | 0.005 | 0.033 | 0.035 | 0.054 | 0.036 | 0.050 | 0.010 | 0.029 | 0.020 | 0.021 |
| t-value | -0.19 | -1.31 | 0.56 | -1.02 | 2.77* | 1.59 | 0. 11 | 2.27* | 1. 12 | -1.22 | 1. 26 | 0.69 | 1.58 |
| **$R^2$(m)** | | | | | | | 0.004 | | | | | | |
| **$R^2$(c)** | | | | | | | 0.278 | | | | | | |

**Note:** Coefficients and standard error are presented for each variable in the baseline model: The modifier lexeme of word length (CL), compound word frequency (CFreq), the head lexeme of frequency (HFreq), the modifier lexeme of frequency (MFreq) and bilabial, labiodental, dental, labioveolar, veolar, postalalveolar, palatal.alveolar (PA), glottal, velar and voiced, with the alveolar factor being removed as a result of its VIF being above 3, indicating moderate co-linearity. 95% CI - = Lower confidence interval; 95% CI+ = Upper confidence Interval and SE = Standard error.

*Significant at the α = .05 level.

+Significant at the. *p* < .10 level.

[a] This model did not converge with the variable of interest. We only included a random subject- and item intercept.

**Table 4.** Linear mixed effects regression results for the baseline model for the head lexeme[a].

| Values | HL | CFreq | MFreq | HFreq | Bilabial | Labiodental | Dental | Labiovelar | postalveolar | PA | Glottal | Velar | Voiced |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *β* | -0.009 | 0.006 | 0.002 | 0.001 | 0.019 | 0.028 | 0.056 | 0.009 | -0.046 | 0.003 | 0.019 | 0.021 | 0.006 |
| SE | 0.004 | 0.004 | 0.004 | 0.004 | 0.006 | 0.009 | 0.032 | 0.007 | 0.016 | 0.022 | 0.009 | 0.008 | 0.005 |
| 2.5% CI | -0.017 | -0.003 | -0.007 | -0.008 | 0.008 | 0.011 | -0.007 | -0.005 | -0.077 | -0.041 | -0.006 | 0.006 | -0.004 |
| 97.5% CI | 0.00003 | 0.011 | 0.011 | 0.009 | 0.030 | 0.046 | 0.190 | 0.023 | 0.029 | 0.046 | 0.036 | 0.036 | 0.016 |
| t-value | -1.96+ | 1.39 | 0.43 | 0.12 | 3.42* | 3.23* | 1.75 | 1.32 | -2.84* | 0.11 | 2.04* | 2.68* | 1.21 |
| **$R^2$(m)** | | | | | | | 0.008 | | | | | | |
| **$R^2$(c)** | | | | | | | 0.245 | | | | | | |

**Note:** Coefficients and standard error are presented for each variable in the baseline model: The modifier lexeme of word length (CL), compound word frequency (CFreq), the head lexeme of frequency (HFreq), the modifier lexeme of frequency (MFreq) and bilabial, labiodental, dental, labioveolar, veolar, postalalveolar, palatal.alveolar (PA), glottal, velar and voiced, with the alveolar factor being removed as a result of its VIF being above 3, indicating moderate co-linearity. 95% CI - = Lower confidence interval; 95% CI+ = Upper confidence Interval and SE = Standard error.

*Significant at the α = .05 level.

+Significant at the. *p* < .10 level.

[a] This model did not converge with the variable of interest. We only included a random subject- and item intercept.

**Table 5.** Linear mixed effects regression results for the 8 variables of interest for the modifier lexeme.

| Values | CFA[a] | CAoA[a] | ST[a] | LMD[a] | MFA[a] | MAoA[a] | HFA[b] | HAoA[a] |
|---|---|---|---|---|---|---|---|---|
| β | -0.011 | 0.014 | -0.006 | -0.010 | -0.011 | 0.016 | -0.004 | 0.008 |
| SE | 0.005 | 0.047 | 0.005 | 0.005 | 0.006 | 0.006 | 0.006 | 0.006 |
| 2.5% CI | -0.021 | 0.005 | -0.016 | -0.020 | -0.021 | 0.004 | -0.015 | -0.003 |
| 97.5% CI | -0.001 | 0.023 | 0.004 | -0.0001 | 0.0003 | 0.028 | 0.007 | 0.020 |
| t-value | -2.16* | 2.98* | -1.23 | -2.17* | -1.91+ | 2.66* | -0.74 | 1. 48 |
| R2(m) | 0.005 | 0.005 | 0.004 | 0.005 | 0.005 | 0.005 | 0.004 | 0.004 |
| R2(c) | 0.278 | 0.278 | 0.278 | 0.278 | 0.278 | 0.278 | 0.278 | 0.278 |

**Note:** Coefficients and standard error are presented for each variable when added separately to the baseline model containing the modifier lexeme of word length (CL), compound word frequency (CFreq), the head lexeme of frequency (HFreq), the modifier lexeme of frequency (MFreq) and initial phoneme onset, which includes bilabial, labiodental, dental, labioveolar, veolar, postalalveolar, palatal.alveolar, glottal, velar and voiced, with the alveolar factor being removed as a result of its VIF being above 3, indicating moderate co-linearity. 95% CI - = Lower confidence interval; 95% CI+ = Upper confidence Interval; SE = Standard error; CFA = Compound word familiarity; CAoA = Compound word Age of acquisition, ST = semantic transparency, LMD = lexeme meaning dominance; MFA = The modifier lexeme of familiarity, MAoA = The modifier lexeme of age of acquisition, HFA = The head lexeme of second lexeme familiarity, HAoA = The head lexeme of age of acquisition.

*Significant at the α = .05 level.

+Significant at the. $p < .10$ level.

[a] This model did not converge. For these analyses, a by-subject random slope was only included for the variable of interest.

[b] This model did not converge with the variable of interest. We only included a random subject- and item intercept.

**Table 6.** Linear mixed effects regression results for the 8 variables of interest for the head lexeme.

| Values | CFA[a] | CAoA[a] | ST[a] | LMD[a] | MFA[a] | MAoA[a] | HFA[b] | HAoA[a] |
|---|---|---|---|---|---|---|---|---|
| β | -0.007 | 0.010 | -0.010 | -0.011 | -0.007 | 0.008 | 0.005 | 0.003 |
| SE | 0.005 | 0.004 | 0.005 | 0.004 | 0.005 | 0.005 | 0.005 | 0.005 |
| 2.5% CI | -0.016 | 0.002 | -0.018 | -0.020 | -0.016 | -0.002 | -0.005 | -0.007 |
| 97.5% CI | 0.002 | 0. 189 | -0.001 | 0.003 | 0.003 | 0.018 | 0.015 | 0.014 |
| t-value | -1.51 | 2.40* | -2.30* | -2.60* | -1.31 | 1.51 | 1.00 | 0.65 |
| R2(m) | 0.008 | 0.009 | 0.008 | 0.009 | 0.008 | 0.008 | 0.008 | 0.008 |
| R2(c) | 0.245 | 0.245 | 0.244 | 0.245 | 0.245 | 0.245 | 0.245 | 0.245 |

**Note:** Coefficients and standard error are presented for each variable when added separately to the baseline model containing the modifier lexeme of word length (CL), compound word frequency (CFreq), the head lexeme of frequency (HFreq), the modifier lexeme of frequency (MFreq) and initial phoneme onset, which includes bilabial, labiodental, dental, labioveolar, veolar, postalalveolar, palatal.alveolar, glottal, velar and voiced, with the alveolar factor being removed as a result of its VIF being above 3, indicating moderate co-linearity. 95% CI - = Lower confidence interval; 95% CI+ = Upper confidence Interval; SE = Standard error; CFA = Compound word familiarity; CAoA = Compound word Age of acquisition, ST = semantic transparency, LMD = lexeme meaning dominance; MFA = The modifier lexeme of familiarity, MAoA = The modifier lexeme of age of acquisition, HFA = The head lexeme of second lexeme familiarity, HAoA = The head lexeme of age of acquisition.

*Significant at the α = .05 level.

+Significant at the. $p < .10$ level.

[a] This model did not converge. For these analyses, a by-subject random slope was only included for the variable of interest.

[b] This model did not converge with the variable of interest. We only included a random subject- and item intercept.