Modality Effects in Vocabulary Acquisition

Merel C. Wolf (merel.wolf@mpi.nl)[1]
Alastair C. Smith (alastair.smith@mpi.nl)[1]
Antje S. Meyer (antje.meyer@mpi.nl)[1]
Caroline F. Rowland (caroline.rowland@mpi.nl)[2]

[1] Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands
[2] ESRC LuCiD Centre & Department of Psychological Sciences,
University of Liverpool, Bedford Street South, L69 7ZA, Liverpool, United Kingdom

Abstract

It is unknown whether modality affects the efficiency with which humans learn novel word forms and their meanings, with previous studies reporting both written and auditory advantages. The current study implements controls whose absence in previous work likely offers explanation for such contradictory findings. In two novel word learning experiments, participants were trained and tested on pseudoword - novel object pairs, with controls on: modality of test, modality of meaning, duration of exposure and transparency of word form. In both experiments word forms were presented in either their written or spoken form, each paired with a pictorial meaning (novel object). Following a 20-minute filler task, participants were tested on their ability to identify the picture-word form pairs on which they were trained. A between subjects design generated four participant groups per experiment 1) written training, written test; 2) written training, spoken test; 3) spoken training, written test; 4) spoken training, spoken test. In Experiment 1 the written stimulus was presented for a time period equal to the duration of the spoken form. Results showed that when the duration of exposure was equal, participants displayed a written training benefit. Given words can be read faster than the time taken for the spoken form to unfold, in Experiment 2 the written form was presented for 300 ms, sufficient time to read the word yet 65% shorter than the duration of the spoken form. No modality effect was observed under these conditions, when exposure to the word form was equivalent. These results demonstrate, at least for proficient readers, that when exposure to the word form is controlled across modalities the efficiency with which word form-meaning associations are learnt does not differ. Our results therefore suggest that, although we typically begin as aural-only word learners, we ultimately converge on developing learning mechanisms that learn equally efficiently from both written and spoken materials.

Keywords: modality effects; word learning; vocabulary acquisition; reading

Introduction

Novel words can be encountered through listening to speech or through reading text. Inherent properties of each modality will have specific processing demands and will pose specific constraints on the learning mechanisms that enable learning in these modalities. It is, however, not yet understood whether these modality-specific demands influence the efficiency of learning in these modalities. The present study aimed at investigating to what extent the modality in which information is presented affects the efficiency of learning novel word form – meaning associations.

The existing literature shows conflicting findings regarding the effect of modality on novel word learning. Concerning word form learning only, benefits have been found in favour of the spoken modality (Bakker, Takashima, Van Hell, Janzen, & McQueen, 2014; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005). Multiple theoretical explanations have been proposed for these observed auditory learning benefits. Firstly, it has been argued that learning from spoken input is more efficient as a result of such mechanisms being developmentally and/or evolutionarily older than those operating on written stimuli (Bakker et al., 2014).

Further, evidence suggests that, relative to the visual modality, in the auditory modality stronger associations develop between sequential events (Penney, 1989) and/or that temporal events are more accurately stored (Glenberg & Jona, 1991). Auditory cortices have been suggested to be more sensitive to sequencing information, due to the sequential nature of auditory information (Frost, Armstrong, Siegelman & Christiansen, 2015).

Cognitive load theory (Sweller, Van Merrienboer & Paas, 1998) also predicts a spoken learning benefit when learning word forms and visual meanings (e.g. a picture or graph) in combination. It argues that cognitive overload is less likely under conditions in which information processing can be divide between the visuo-spatial sketchpad and phonological loop (Baddeley, 1992), compared to conditions in which all information must be processed within the same modality and thus by the same cognitive resources.

In contrast to the above, a written advantage has also been observed particularly when word forms are learned in conjunction with their meanings, (Balass, Nelson, & Perfetti, 2010; Nelson, McEvoy, & Schreiber, 2004; Van der Ven, Takashima, Segers, & Verhoeven, 2015). Multiple theories have also been proposed in explanation for these findings. It is argued that when reading (novel words) phonological representations are automatically activated alongside orthographical representations, therefore, two separate representations of the word form are stored. However, on exposure to the spoken word form, automatic activation of its orthographic form is less likely (Perfetti, Bell & Delaney, 1988; Paivio, 1991). Further, the spoken modality is fleeting by nature, posing additional demands on attention and working memory capacity. Reading allows rereading and processing at one's own pace and this flexibility leads to greater availability of memory and attentional cognitive resources for learning (Van der Ven, 2015).

Alternatively, in contrast to the above findings it remains possible that learning mechanisms operating on written and spoken stimuli are equally efficient and instead observed contradictory effects result from modality specific biases in the experimental design. Although typically, prior to literacy, word learning is only possible via the auditory modality, it is feasible that proficient readers develop learning mechanisms that overcome modality specific constraints such that learning occurs equally effectively in both modalities.

Previous studies, that have reported modality effects, have potentially generated contradictory findings due to an absence of one or more of the following controls. First, exposure duration was not controlled in studies that found a written learning advantage (Balass et al., 2010; Nelson et al., 2006; Van der Ven et al., 2015). People were given unlimited time with the spoken and written materials, but the exact exposure time was not measured. Participants thus might have exposed themselves more to materials in one modality, evoking a learning effect that does not result from a more efficient modality specific learning mechanism but simply due to a mechanism having greater exposure to the stimulus.

Second, in all studies that found a written benefit the test was presented in a written form (Balass et al., 2010; Nelson et al., 2006; Van der Ven et al., 2015); likewise, some studies that found a spoken benefit performed only a spoken test (Van der Elst et al., 2005). According to Tulving and Thomas's (1973) encoding specificity principle, recall is enhanced if the conditions during retrieval match the conditions during learning. Thus, such modality effects observed in these studies might be evoked by encoding specificity rather than by differences in the efficiency of the spoken and written learning mechanisms. Similarly, studies examining learning of word form-meaning associations only used written meanings. Thus, the congruency of the format between written word forms and written word meanings potentially benefits learning in the written modality.

Fourth, many previous studies have used explicit learning tasks (Balass et al., 2010; Nelson et al., 2006; Van der Ven et al., 2015). Therefore, in such studies, it is difficult to exclude the possibility that observed modality effects do not result from modality-specific conscious learning strategies, such as repeating heard words or rereading written words, rather than differences in the efficiency of modality specific cognitive mechanisms.

Finally, many previous studies (Bakker et al., 2014; Balass et al., 2010; Nelson et al., 2006; Van der Ven et al., 2015; Van der Elst et al., 2005) do not control for cross-modal orthographical and phonological transparency. Therefore, any learning benefit observed may not result from differences in the efficiency of learning mechanisms but instead may result from it being easier to accurately transform the phonological form to the orthographic or vice versa.

In order to gain an understanding of modality effects on word learning it is first necessary to control for each of these potential confounds. The present study aims to do precisely this, controlling for the many confounds that have potentially generated observed modality effects that do not result from difference in efficiency of the spoken and written learning mechanisms.

In two experiments, participants learned 24 Dutch-like, fully transparent pseudowords and pictorial meanings. After a short period of consolidation, participants were tested on their knowledge of the learned word forms and meanings. A between-subjects design generated four participant groups per experiment 1) written training, written test; 2) written training, spoken test; 3) spoken training, written test; 4) spoken training, spoken test. In addition, non-verbal IQ, vocabulary and reading tasks were administered to control for differences across groups. In Experiment 1 written word forms were presented for a time period equal to the duration of the spoken form. In Experiment 2, to control for the fact that a written word can be read quicker than its spoken form takes to unfold, the written stimulus was presented only for the period necessary to read the written stimulus.

Experiment 1

Methods

Participants 60 participants (*M* = 22.96 years, *SD* = 2.53; 46 female) were recruited. All participants were right-handed, with no language, sight or hearing disorders. Participants earned €10 for participating.

Design The two between-subjects factors were modality during training and modality during testing. Words could be learned in either modality and also testing could occur in two modalities. There were therefore four between-subjects conditions. Participants were randomly assigned to a condition.

Materials Twenty-four orthographically and phonologically transparent Dutch pseudowords were created using Wuggy (Keuleers & Brysbaert, 2010). The words had a Levenshtein's Distance (Levenshtein, 1966) of above three to avoid confusability. Pilot studies ensured the words were not reminiscent of existing Dutch words. The words varied between five and nine letters and four and eight phonemes and graphemes. Speech duration of the words varied between 664 and 993 ms.

In addition twenty-four pictures of unknown objects from The Novel Object and Unusual Name (NOUN) Database were used (Horst & Hout, 2016). The pictures were not visually similar to each other. To limit item-specific effects, for each group of four participants the pictures were randomly assigned to one of the word forms.

Procedure Participants were trained and tested on the same day. First the training phase was administered. The experiment was designed to minimize opportunities for participants to utilize explicit learning strategies. For this reason no explicit instruction to learn the picture–word form pairs or indication of a later test was provided, images and word forms were presented briefly and in rapid succession, and both auditory and visual masks immediately followed presentation of the word form. In each trial (Figure 1), participants saw a fixation cross (250 ms), a picture (1000 ms), then again saw a fixation cross (250 ms), either heard the word or read the word depending on the condition, and then heard a auditory mask in the form of a continuous tone and saw a visual mask in the form of a grey diamond (500 ms). The exposure to the word form varied for each word: the written word was presented for the speech duration of that specific word (M = 863 ms, SD = 97 ms). The next word in the training sequence always had a Levenshtein's distance above three and a different onset. Each training trial was repeated seven times in a blocked, semi-randomized order. To ensure attention during the training phase, eight pictures of familiar known objects (e.g. a bus) were shown in-between the trials and participants had to press a button as soon as they saw one of these familiar objects. Participants were instructed to pay attention to the pictures and words and press a button if they saw one of the eight familiar objects, but critically were not explicitly told to learn the word form – picture pairs.

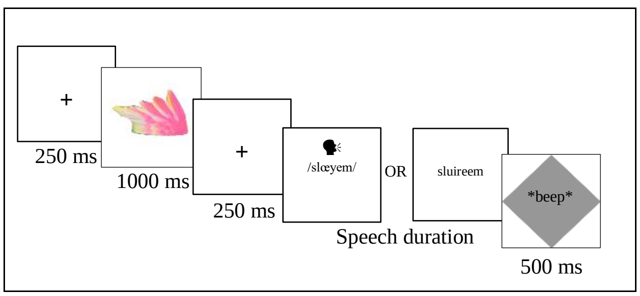

Figure 1: Experimental procedure of a training trial

The training phase was followed by a filler task. This purely visual, nonverbal IQ-task lasted for 20 minutes (Raven's progressive matrices, 1965). Then, in the test phase, participants performed a subsequent matching task. Participants saw a fixation cross (250 ms), a picture (1000 ms), then again a fixation cross (250 ms), heard or saw the word depending on the condition, and had to decide within 2 seconds whether the picture and word matched what they had learned by using a button box. The written words were again presented for a time period equal to the speech duration of that particular word. Each word was presented twice: once with the correct picture and once with a foil picture (i.e., a different picture presented in the training phase). There were several constraints regarding the relationship between the foil picture and the target word form. The corresponding learned word form of the foil picture did not share the onset of the target word form and possessed a Levenshtein's distance of above four. Regarding the order of the trials, the corresponding word form of the next (foil) picture could not be one of the previous ten word forms. Also, half of the target words were first shown with the correct picture before they appeared with a foil picture and vice versa. Participant's ability to identify both matching and mismatch picture-word form pairs was recorded. Then, several individual difference tests were administered, including word reading, pseudoword reading (Van den Bos, Spelberg, Scheepsma & de Vries, 1994; Brus & Voeten, 1973) and vocabulary (Dunn, Dunn, & Schlichting, 2005).

Results

Violin plots depicting, per condition, the proportion of picture–word form pairs that were correctly identified as a match or mismatch can be found in Figure 2.

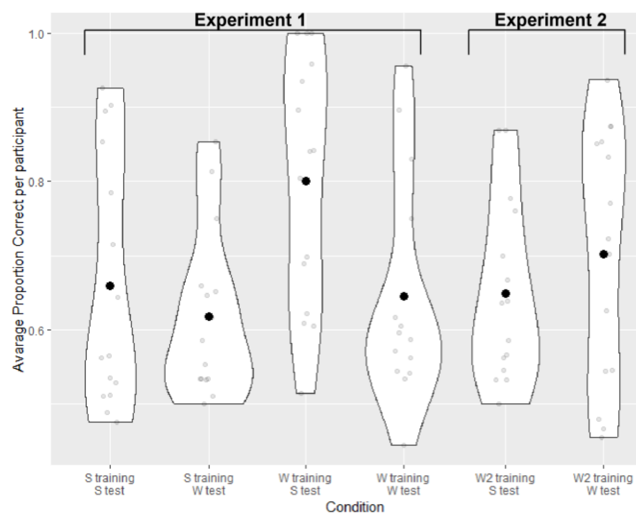

Figure 2: Proportion of correctly identified matching and mismatching picture-word form pairings per participant

A mixed effects logistic regression model (lme4 package: Bates, Maechler, Bolker, & Walker, 2015) using R (R Development Core Team, 2008) was constructed with response on test (match or mismatch) as the dependent variable, i.e. whether a participant recorded the corresponding image and word form pair as matching or mismatching. Model structure was compatible with the conventions of standard signal detection analysis and was consistent with current best practice (e.g. Jacobs, Dell, Benjamin, & Bannard, 2016; Zormpa, Brehm, Hoedemaker & Meyer, 2019). The model included fixed effects of trial type (whether the trial was a match or mismatch), training modality (written or spoken) and test modality (written or spoken), in addition to their interactions. The full random effect structure was also included in the model with random intercepts and slopes by item for trial type, training modality and test modality, and random intercepts and slopes by participant for trial type.

Model results revealed a main effect of trial type, showing participants displayed sensitivity to trained versus untrained picture–word pairs, providing a match more frequently when presented with the picture – word pairs on which they were trained (estimate = -1.03, *SE* = 0.16, *z* = -6.35, *p* < .001). The interaction between trial type and training modality was also significant (estimate = 0.29, *SE* = 0.13, *z* = 2.20, *p* = .03) with participants in trained on written word forms displaying greater sensitivity in identification of trained vs. untrained picture – word for pairs. Finally, a significant interaction between trial type and test modality was also observed (estimate = -0.37, *SE* = 0.13, *z* = -2.78, *p* = 0.006) with participants displaying greater sensitivity when tested on spoken word forms.

Conclusion

Experiment 1 results show that when controlling for exposure time by providing equal exposure duration in both modalities, learning from written materials is greater. One explanation for this might be differences between modalities in the speed with which the full word form can be accessed from the stimulus. The speech duration of the word forms was between 664 and 993 ms, and thus, the written words were presented for a duration of between 664 and 993 ms, depending on the word. First pass single word reading is however much faster than the reading time provided in Experiment 1. Literature using lexical decision or naming tasks show that bisyllabic word can be read at between 525-610 ms, and pseudowords between 575 and 650 ms (Brunswick, McCrory, Price, Frith, & Frith, 1999; De Groot & Nas, 1991; Schilling, Rayner, & Chumbley, 1998; Weekes, 1997). However, these estimates include time necessary to make a decision and speech planning. Studies using ERP and eye-gaze measures, which give a more accurate estimate of reading times, show that frequent, known words can be read around 150 ms and infrequent words within 200-250 ms (Rayner, Pollatsek, Ashby, & Clifton Jr, 2012; Schilling et al., 1998; Sereno, Rayner, & Posner, 1998). This means that, although exposure time to the written and spoken stimuli was equal in Experiment 1, people had more time with the full word form in the written condition.

Experiment 2 tested whether the modality effects found in Experiment 1 would hold if exposure to written and spoken materials was equivalent, taking into account that written information is presented instantaneously and that reading is faster than listening to speech. Literature has shown that people need slightly longer to read infrequent words (200-250 ms) than frequent words (150 ms). Pseudowords are thus likely to be read slightly slower. Therefore, in Experiment 2, the written exposure time was set at 300 ms for all 24 words, which is a written exposure time reduction of 65% on average relative to Experiment 1.

Experiment 2

Methods

Participants 30 participants (*M* = 23.02 years, *SD* = 2.40; 26 female), all right-handed, with no language, sight or hearing disorders participated in this experiment. Participants earned €10,- for participating.

Design Experiment 2 only concerned written modality learning. Testing occurred in both modalities, creating two conditions. Participants were randomly assigned to a condition.

Materials The materials were the same as in Experiment 1.

Procedure The procedure was similar to that of Experiment 1, except for the training phase. In the training phase, the written word was now presented for 300 ms rather than the speech duration of that specific word. This reduced the total duration of the training phase by 560 ms. To ensure that this shortening of the trial did not affect learning, in each trial the first fixation cross was elongated from 250 to 530 ms and the mask at the end of a trial was elongated from 500 to 780 ms. After training participants again performed a non-verbal IQ test, followed by the picture-word form matching task and the individual difference measures.

In addition, to test that 300 ms was sufficient time for participants to read the word-forms, a simple retyping task was added to test whether participants could read 120 additional Dutch pseudowords equally well when presented for either 300 ms or 860 ms (the mean written exposure time of Experiment 1). This retyping task was only administered to the participants in the written training condition.

Results

One participant from Experiment 2 had to be removed, because no buttons were pressed during the matching task. Violin plots of the accuracy data can be found in Figure 2. Four one-way ANOVA's indicated that the six groups (four from Experiment 1 and two from Experiment 2) did not differ regarding average general IQ ($F(5,83) = 0.46$, $p = .81$), vocabulary ($F(5,83) = 0.64$, $p = .67$), word reading ($F(5,83) = 0.69$, $p = .63$) or pseudoword reading ability ($F(5,83) = 0.67$, $p = .65$).

To analyse performance on the retyping task, a frequentist mixed-effect logistic regression model was applied using R package lmer (lme4 package: Bates, Maechler, Bolker, & Walker, 2015) with retyping accuracy as dependent variable, and word length and exposure time (300 or 860 ms) as independent variables, plus a random intercept by participant and word. This analysis showed no difference in accuracy of retyping after a 300 or 860 ms exposure (estimate = 0.54 *SE* = 1.81, *z* = 0.29, *p* = .77).

The mixed-effects logistic regression model used to analyse results in Experiment 1, was extended to analyse results of both experiments, with modality at training now possessing three levels: spoken training in Experiment 1, written training in Experiment 1 where written exposure time was equal to spoken exposure time, and written training in Experiment 2 where written exposure time was reduced to 300 ms. The bias effects of modality at training and test on hits and false alarms are illustrated in Figure 2.

Analyses revealed a significant main effect of trial type with participants more likely to produce a match response when trials included the picture – word form pairs on which they were trained (estimate = -0.73, *SE* = 0.19, *z* = -3.89, $p < 0.001$). The interaction between trial type and training modality was not significant when comparing the reduced written training condition (Experiment 2) to that of the spoken training condition (estimate = -0.24, *SE* = 0.25, *z* = -0.96, *p* = 0.34) indicating that sensitivity of participants did not differ significantly between groups. Similarly, the interaction between trial type and training modality was not significant when comparing the reduced written training condition (Experiment 2) to the longer written training condition (Experiment 1) (estimate = -0.38, *SE* = 0.25, *z* = -1.52, *p* = 0.13). The three-way interaction between training modality, test modality and trial type was significant when comparing the two written conditions (estimate = -0.73, *SE* = 0.25, *z* = -2.90, *p* = 0.004). The three-way interaction was not significant (estimate = -0.43, *SE* = 0.25, *z* = -1.69, *p* = 0.09) when comparing the longer written training condition (Experiment 1) to the spoken training condition or the reduced written training condition (Experiment 2) to the spoken training condition (estimate = 0.31, *SE* = 0.25, *z* = 1.25, *p* = 0.21). Thus, participants trained in the longer written condition (Experiment 1) displayed greater sensitivity during the spoken test than participants trained in the shorter written training condition (Experiment 2) or spoken training condition.

Conclusion

Experiment 2 aimed to investigate whether the written modality benefit found in Experiment 1 resulted from participants having more time with the word form in the written condition, due to the fact that it takes longer for a spoken word to unfold than to read its written from. By reducing written word exposure to 300 ms per word, we controlled for this inherent advantage of the written modality. Results showed that when the exposure time to the written materials was reduced, learning in the written condition did not differ from that in the spoken condition. Further, this was not a result of participants having insufficient time to read the written form as participants did not differ in their ability to retype written pseudowords when they were presented for 300 ms or 860 ms.

General Discussion

This study aimed to test whether modality specific learning mechanisms, engaged when learning novel picture–pseudoword form pairings, are more effective when words are presented in their written or spoken form. This study is the first to test for such effects of modality while controlling for the following factors, which potentially give rise to modality effects independent of differences in the efficiency of modality specific learning mechanisms: 1) differences in orthographic and phonological transparency, 2) congruence in modality of word form and word meaning, 3) duration of exposure, 4) engagement of explicit learning strategies, 5) congruence in modality of training and modality of test.

Our results showed that when the duration of written and spoken exposure is equal (the written stimulus is presented for a time period equal to the duration of the spoken word), participants' accuracy in identifying picture-word form pairs is greater when trained on written word-forms. This finding replicates earlier findings of a written learning benefit when learning word forms and their meanings (Balass et al., 2010, Nelson et al., 2004, Van der Ven et al., 2015).

However, Experiment 2 shows that the written learning benefit disappears when controlling for the fact that the time required to read a word in its written form is shorter than the time required for its spoken form to unfold. Our results demonstrate that once controlling for this property of reading there is no additional advantage in learning word form – picture associations when words are presented in their written rather than spoken form.

Our conclusions are therefore at odds with previous studies that argue for differences in the efficiency of modality specific learning mechanisms. Based on the results produced by this study we believe such findings are likely driven by an absence of one or more of the confounds listed above (see list 1-5), which alone may generate such observed modality effects.

Bakker et al, (2014), one of few studies to train and test participants in both modalities, provides evidence that auditory benefits of learning novel word forms emerge only at longer periods of consolidation. Within their study phoneme and letter monitoring tasks were used to probe lexical integration of novel word forms after 24 hrs and 8 days. It is feasible therefore that the findings within the current study are limited to short-term episodic memory. This can be tested in a follow up study by extending the current paradigm to include tests of lexical integration at longer periods of consolidation.

Unexpectedly, our results did not produce a modality congruency effect as predicted by Tulving and Thomas's (1973) encoding specificity principle, in that the experimental groups for which the test modality was the same as the training modality, did not show superior performance. Paradoxically, the written benefit observed in Experiment 1 was mainly driven by the written learning spoken test group. However, we believe this to be caused by the perceived erratic response window in the written test condition. Participants were required to respond within 2 seconds plus the speech duration of the written word. Because they did not hear the word, the response time was therefore difficult to predict. This conclusion is supported by participant's performance on the same task in Experiment 2, when participants were habituated in the written training phase to a fixed exposure time which did not appear to result in a decrease in performance on the written test.

Still, our experiments do not provide evidence for Tulving and Thomas's (1973) encoding specificity principle, since participants in all cross-modal conditions were consistently able to recognize words in a modality in which they had not seen the word form before. Further, no interaction was observed between the written reduced training condition of Experiment 2 and the spoken test condition of Experiment 1, indicating that when participants have equivalent exposure to either the written or spoken word form in training, their ability to recognise the novel word form in the alternative, unseen modality does not differ.

Within the current study, attempts were made to limit strategic cross-modal encoding: no explicit instruction to learn the materials was provided, participants were trained in a single modality, stimuli were presented rapidly, and visual and auditory masks immediately followed the presentation of the word form. Thus, our results suggest that proficient readers, such as those tested in our study, automatically rapidly recode novel word forms into both their phonological form when presented with written stimulus (Perfetti et al., 1988) and their orthographic form when presented with an auditory stimulus.

Our findings also do not support a developmental and/or evolutionary advantage for learning from spoken materials. It appears that even though the ability to learn from written materials has developed later in human's lives and their evolution as a species, this ability is sufficiently developed in adult proficient readers to perform equally effectively.

This study set out to test for modality effects on novel word learning. Specifically it tested for differences in the efficiency of modality specific mechanisms engaged when learning novel object - pseudoword pairs, from either spoken or written stimuli. Results showed a written benefit when equal exposure time was provided. However, once we controlled for the fact that reading allows faster access to the full word form than listening to speech, no modality effect was observed. This suggests that modality specific learning mechanisms operating on spoken or written stimuli were equally efficient. Given that we typically begin learning words from auditory input only, the findings of the present study indicate that once we become proficient readers, the cognitive system converges on learning equally efficiently from both modalities.

References

Bakker, I., Takashima, A., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language, 73,* 116-130.

Balass, M., Nelson, J. R., & Perfetti, C. A. (2010). Word learning: An ERP investigation of word experience effects on recognition and word processing. Contemporary *Educational Psychology, 35*(2), 126-140.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz? *Brain, 122*(10), 1901-1917.

Brus, B. T., & Voeten, M. (1973). Een-minuut test. Vorm A en B. *Verantwoording en handleiding*.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*(1), 90-123.

Dunn, L. M., Dunn, L. M., & Schlichting, J. E. P. T. (2005). *Peabody picture vocabulary test-III-NL*. Amsterdam, Netherlands: Harcourt Test Publishers.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cognitive Sciences, 19*(3), 117-125.

Glenberg, A. M., & Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects. *Memory & Cognition, 19*(5), 514-522.

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods, 48*(4), 1393-1409.

Jacobs, C. L., Dell, G. S., Benjamin, A. S., & Bannard, C. (2016). Part and whole linguistic experience affect recognition memory for multiword sequences. *Journal of Memory and Language, 87, 38-58.*

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627-633.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10*(8), 707-710.

Nelson, J. R., Balass, M., & Perfetti, C. A. (2006). Differences between written and spoken input in learning new words. *Written Language & Literacy, 8*(2), 25-44.

Paivio, A. (1991). Dual Coding Theory - Retrospect and Current Status. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 45*(3), 255-287.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*(4), 398-422.

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27*(1), 59-70.

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Raven, J. (1965). *Advanced Progressive Matrices. Sets I and II.* London: HK Lewis & Co.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of Reading, 2nd ed.* New York, NY, US: Psychology Press.

Schilling, H. H., 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.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: evidence from eye movements and event-related potentials. *Neuroreport*, *9*(10), 2195-2200.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Tulving, E., & Thomson, D. M. (1973). Encoding Specificity and Retrieval Processes in Episodic Memory. *Psychological Review, 80*(5), 352-373.

Van den Bos, K., Spelberg, H., Scheepsma, A., & De Vries, J. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden. Verantwoording, handleiding, diagnostiek en behandeling.* Nijmegen: Berkhout.

Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2005). Rey's verbal learning test: normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society, 11*(3), 290-302.

Van der Ven, F., Takashima, A., Segers, E., & Verhoeven, L. (2015). Learning word meanings: overnight integration and study modality effects. *PLoS One, 10*(5), e0124926.

Weekes, B. S. (1997). Differential Effects of Number of Letters on Word and Nonword Naming Latency. *The Quarterly Journal of Experimental Psychology Section A, 50*(2), 439-456.

Zormpa, E., Brehm, L. E., Hoedemaker, R. S., & Meyer, A. S. (2019). The production effect and the generation effect improve memory in picture naming. *Memory, 27*(3), 340-352.