Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Individual Differences and Lexical Learning: Links to memory for faces, things, and words

Permalink

https://escholarship.org/uc/item/07t0017s

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN

1069-7977

Authors

Shaw, Ashlee
Demos, Alexander
Arthur, Dana
et al.

Publication Date

2011

Peer reviewed

Individual differences and lexical learning:
Links to memory for faces, things and words

Ashlee Shaw (ashlee.shaw@uconn.edu)[1,2]
Alexander P. Demos (alexander.demos@uconn.edu)[1,2]
Dana Arthur (dana.arthur@uconn.edu)[3]
James S. Magnuson (james.magnuson@uconn.edu)[1,2]
[1]Department of Psychology, University of Connecticut, Storrs, CT 06250
[2]Haskins Laboratories, 300 George St., New Haven, CT 06511
[3]Communication Disorders, University of Connecticut, Storrs, CT 06250

Abstract

The Lexical Quality Hypothesis (Perfetti & Hart, 2002) suggests that the difficulties exhibited by poor readers cascade from deficient (impoverished, fuzzy) representations of phonological, semantic, and orthographic dimensions in lexical memory. If so, readers, even as adults, should vary in their ability to acquire new lexical representations. In our study, we examine the role of cross-modal (visual to phonological) associations in lexical learning. By pairing an artificial lexicon with novel objects, we aim to see whether learning implicit associations between new words and visual features of novel objects can be predicted by participants' performance in a number of visual and language-related assessments. We report intriguing preliminary results suggesting new relationships between recognition memory and ability for language learning and processing.

Keywords: Language, learning, face recognition.

Introduction

Perfetti and Hart's (2002) Lexical Quality Hypothesis (LQH) posits that most difficulties with reading comprehension can be linked causally to difficulties with the strength and richness of an individual's word-level knowledge. A high quality lexical representation incorporates detailed orthographic, semantic, and phonological information. The theory posits that the stronger and more specific the information contained within a lexical representation is, the more efficiently that word can be accessed during reading. According to this hypothesis, less skilled readers possess weak or unclear lexical representations that are not optimal for efficient access, which cascades to problems with comprehension (Perfetti & Hart, 2002).

The "triangle model" of reading (e.g., Harm & Seidenberg, 1999, 2004; Figure 1) provides a mechanistic analog to the LQH. The triangle refers to initially learned connectivity patterns between phonological and semantic representations, and later learning of phonological-orthographic and semantic-orthographic mappings. A model trained with a typical training regimen learns, for example, to rely most heavily on the phonological→orthographic pathways for regular sound-spelling patterns and more heavily on phonological→semantic→orthographic pathways for irregular sound-spelling patterns. To the degree that phonological or semantic representations or phonological-semantic pathways are noisy or weak prior to orthographic learning, the model will be at a severe disadvantage when orthographic training begins.

This begs the question: what kinds of individual differences in cognitive abilities might lead to noisy or weak representations or pathways and hence to low lexical quality? While variation in linguistic ability is a logical candidate, other factors might contribute, such as memory ability, associative learning ability, or the ability to map information across modalities, such as from objects to names, or from names to print.

One way to examine acquisition of phonological-to-semantic connections is by using a spoken artificial lexicon (Magnuson, Tanenhaus, Aslin & Dahan, 2003). Utilizing an artificial lexicon allows us to tightly control properties of linguistic and visual materials and ensure that each participant has no prior experience with the stimuli, minimizing potential differences in lexical dimensions (e.g., word frequency) and preexisting semantic associations. The paradigm also allows us to observe any learning effects from the very beginning of the experiment. Thus, this paradigm allows us to put readers who vary in reading ability on a maximally similar level with regard to prior knowledge and language experience with our experimental items.

Recently, this paradigm has been applied to the study of individual differences across a wide range of reading skill. Magnuson, Kukona, Braze, Johns, Van Dyke, Tabor, Mencl, Pugh, and Shankweiler (2010) found that performance on standard assessments like rapid auditory naming predicted the degree to which low-literacy adults exhibit lexical competition effects and how sensitive they are to coarticulation. However, while that project included dozens of language measures, it included only a few standardized assessments of non-linguistic abilities. This leads to complementary questions we address here:

- What sorts of individual differences will we observe in linguistic and non-linguistic abilities in a typical college sample (rather than the low-literacy adults from Magnuson et al., 2010)?
- Will those differences be compatible with the premises of the Lexical Quality Hypothesis (that is, will participants

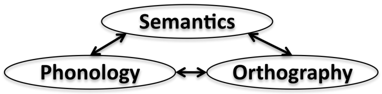

Figure 1: Simple schematic of the triangle model of reading.

on the low end of linguistic ability similarly lag their peers in learning novel words)?

- Alternatively, or perhaps in addition, might performance in learning new words be more strongly associated with simple learning (recognition memory) across domains (faces, objects, spoken words)?

We began our line of questioning by exploring the relationship between semantics and phonology. In our experiment, we examined whether performance scores on standardized tests of language ability or visual and language-related memory tasks could predict readers' ability to link new words to concrete visual objects. From the basis of the Lexical Quality Hypothesis, we predicted that language ability should be closely related to artificial lexicon learning. Our design also allows us to ask whether such differences are specific to language, or might apply more generally across domains.

Methods

Participants

Forty-six University of Connecticut undergraduates were participants in the experiment. All participants were native, monolingual English speakers who reported normal hearing and normal or corrected-to-normal vision.

Apparatus and Materials

Assessments Five assessments were used to measure individuals' abilities in linguistic and nonlinguistic domains. These included tests of verbal working memory (the Reading Span Task [RST] of van den Noort et al., 2008) and word reading efficiency, both of real words and pseudowords (Test of Word Reading Efficiency [TOWRE], Torgeson, Wagner & Rashotte, 1997). We also administered face, object, and spoken word recognition (old/new) tasks of our own construction.

The first assessment was the Reading Span Task (RST; Daneman and Carpenter, 1980; van den Noort et al, 2008), which is a measure of verbal working memory. In the RST, participants read multiple sets of 2-6 sentences aloud, with sentence lengths of approximately 13-16 words. Each sentence ends in a different word. After each group, participants are asked to recall the final words of each sentence in the group. Participants were tested on a total of 60 sentences (see van den Noort et al. for details).

The second assessment administered was the Test of Word Reading Efficiency (TOWRE; Torgeson, Wagner & Rashotte, 1997), which tested participants' word-level reading skills. This timed measure, normed for participants up to 24 years of age, quickly assesses the speed and accuracy of decoding and word recognition. It consists of two subtests. In the Sight Word Efficiency (SWE) subtest, the participant is presented with a list of printed real words and instructed to read aloud as many as possible in 45 seconds. Words in this subtest are arranged in order of decreasing frequency and increasing length. In the Phonetic Decoding Efficiency (PDE) subtest, the participant is presented with a list of pronounceable pseudowords and asked to decode aloud as many as possible in 45 seconds. The pseudowords in this list represent a variety of grapheme-phoneme correspondences and increase in difficulty as the test progresses. Thus, both subtests are designed to increase in difficulty while taxing the participant with added time pressure. Further, from the point of view of the Lexical Quality Hypothesis, the SWE subtest of the TOWRE should shed light on participants' ability to quickly access pre-existing lexical representations.

For the face recognition task, the stimuli were 50 faces taken from Nestor and Tarr (2008), which were approximately balanced in terms of gender and race. During the exposure phase, participants were shown 25 faces for duration of 300 ms each, with a 300 ms inter-stimulus interval. During testing, participants were shown a total of 50 faces, and pressed a key to indicate whether s/he saw the face during exposure. Twenty-four of the faces (12 old, 12 new) during testing were presented in an alternate orientation (i.e., with a left- or right-facing profile of either 30, 45, or 60 degrees).

The object recognition task included 150 realistically-rendered images of objects from the Tarr Object Databank (images courtesy of Michael J. Tarr, Carnegie Mellon University, http://www.tarrlab.org). Roughly equal numbers of objects were selected from 12 taxonomic categories (~12 from each), and were judged by the experimenters to be roughly similar in visual salience. During the exposure phase, participants were shown 75 objects for a duration of 300 ms each, with a 300 ms inter-stimulus interval. During testing, participants were shown a total of 150 images, and pressed a key to indicate whether s/he saw the object during exposure. Half of the objects during testing were presented in an "alternate" orientation (rotated 90- 180 degrees). Thirty-eight of the alternate orientations were of old objects, and 37 were new.

Finally, the old/new spoken word recognition task was constructed as follows: a total of 152 spoken words were recorded by two female speakers (words were 1-7 syllables; average syllables= 2.4). Each of the speakers spoke half of the items for both the old and new sets (76 items total per speaker). Old items were categorized into a "same" or "different" condition -- i.e., the speaker during the exposure phase either was or was not the same speaker during recognition testing. The instructions made clear that a word should be considered "old" even if the voice were not the same. During exposure, participants listened to 76 spoken words (300 ms inter-stimulus interval). During testing, participants heard 152 words and were instructed to press a key indicating whether the spoken word was heard during exposure or not.

Note that space limitations preclude us from presenting results from these old/new tasks in terms of altered orientation or voice. We will simply report d' performance collapsing across these factors.

Artificial lexicon experiment The primary task was to learn the names of nine mushrooms. The mushrooms varied

Table 1: Illustration of visual feature-syllable pairings. C_1 = cap 1, S_1 = stem 1, C_2 = cap 2, etc.

	Correlated condition			Uncorrelated condition		
Feature and artificial lexical item pairings	C_1S_1:pile	C_1S_2:piva	C_1S_3:pisae	C_1S_1:pile	C_1S_2:dova	C_1S_3:gusae
	C_2S_1:dole	C_2S_2:dova	C_2S_3:dosae	C_2S_1:guva	C_2S_2:pisae	C_2S_3:dole
	C_3S_1:gule	C_3S_2:guva	C_3S_3:gusae	C_3S_1:dosae	C_3S_2:gule	C_3S_3:piva

in two visual dimensions: they had one of three caps and one of three stems (see Figure 2). Each mushroom had a two-syllable name, such as /pile/ ("pea-lay"). The names were combinations of three possible first syllables (/pi/ [pea], /do/ [dough], /gu/ [goo]) and three possible second syllables (/le/ [lay], /va/ [vah], /sae/ [as in "sat"]). The relationship between visual and phonological features was manipulated between participants. For participants in the "correlated name" condition (n=25), the syllables mapped directly onto visual properties of the mushrooms, such that the first syllable named the cap and the second named the stem (thus, the name of any mushroom with a particular cap would begin with the same syllable, and the name of any mushroom with a particular stem would have the same second syllable). In the "uncorrelated" condition (n=21), visual and phonological features were completely uncorrelated, such that mushrooms with the same cap or stem had no phonological overlap in that dimension, and mushrooms with the same first or second syllable had no visual overlap in that dimension. Table 1 lists the specific feature-name pairings for each condition.

Procedure

The testing session began with the assessments and the exposure phases of the old/new tasks. This was followed by the artificial lexicon experiment and then the test phases of the old/new tasks.

Participants were assigned randomly to the experimental conditions. Participants were not informed about possible correlations in the materials in either condition. They were simply told to learn the names of the objects, in a 2-alternative forced choice task.

Phonological stimuli were presented auditorally, in the form of instructions such as “Find pile.” Participants responded by clicking on one of the mushrooms. Initially, they just had to guess. If they clicked on the incorrect item, they heard an instruction to "try again." When the participant clicked on the correct item, the incorrect item disappeared, and they heard feedback like "that's right, that's the pile," and then the trial ended. To begin the next trial, participants clicked on a fixation cross in the center of the computer screen. Every 24 trials, a progress report was displayed on the screen, telling the participant his/her percentage correct over the preceding 24 trials, and offering them an opportunity to take a break. Experimental blocks consisted of 72 trials; over the course of a block, participants were tested on each possible stimulus pairing. Trial order was pseudo-randomized in each block so that each stimulus type was distributed equally over the block. There were 5 blocks, for a total of 360 trials.

Figure 2: Example mushrooms. There are three possible caps and three possible stems. Among these examples, the first and second have the same cap, and the first and third have the same stem. No others overlap in stem or cap.

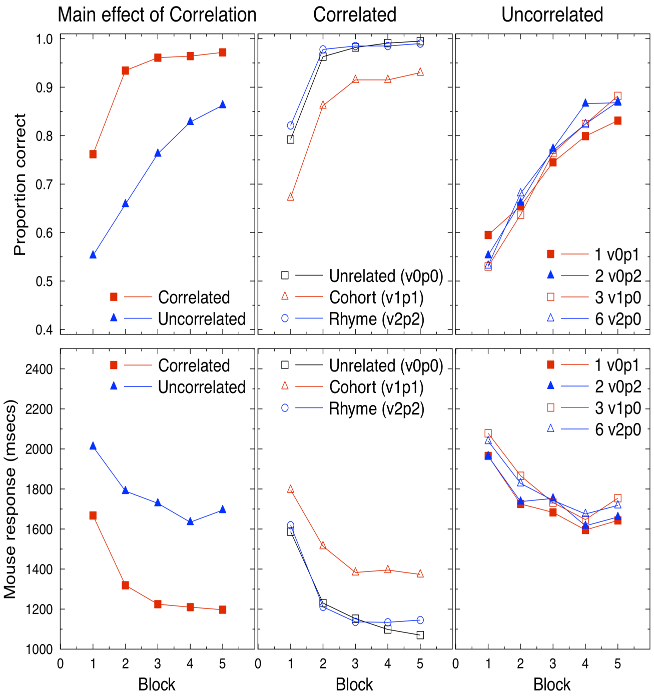

Figure 3. Accuracy (top) and mouse-click reaction time (bottom). Left-most panels compare accuracy and RT for correlated and uncorrelated conditions, collapsing over item types. The center and right panels show correlated and uncorrelated condition results by item type. Labels such as v0p0 indicate overlap in visual (v) and phonological (p) dimensions; v0 = no visual overlap, v1 = same cap, v2 = same stem; p0 = no phonological overlap, p1 = same first syllable, p2 = same second syllable. As described in the text, identical similarity relations cannot occur in correlated and uncorrelated conditions.

Results

Artificial lexicon experiment

Figure 2 shows accuracy and reaction times (RT to click on the target item with the computer mouse only for correct trials) for both experimental conditions across all five trial blocks. Unsurprisingly, participants showed higher accuracy and faster RTs in the correlated condition over the uncorrelated condition for all five blocks. An ANOVA on accuracy by correlation condition and block (collapsing across stimulus type) revealed reliable main effects of correlation condition (correlated = 0.92, uncorrelated = 0.73; $F(1,45)=63.2$, $p<0.001$) and block ($F(4,180)=122.1$, $p<0.001$), as well as a significant interaction of correlation condition and block ($F(4,180)=11.8$, $p<0.001$). The interaction follows from the earlier plateau in the correlated condition. For RT, there were significant main effects of correlation condition (correlated=1323 msecs, uncorrelated =1772 msecs, $F(1,45)=20.4$, $p<0.001$) and block ($F(4,180)=23.4$, $p<0.001$), confirming the trends apparent in Figure 2. The interaction was not significant.

Now let's consider effects of stimulus type within each correlation condition. In the correlated condition, an ANOVA on accuracy revealed reliable main effects of block ($F(4,88)=82.7$, $p<0.001$) and stimulus type ($F(2,44)=16.7$, $p<0.001$), and a reliable interaction ($F(8,176) = 2.1$, $p<0.05$). In the interest of space, we will not unpack all of these in detail, but will simply note that cohort trials (where items shared caps and first syllables) were reliably less accurate than rhyme or unrelated trials, which did not differ from each other. The interaction of block and condition followed from reliable differences between rhyme and unrelated conditions in early blocks that disappeared by block 3. An ANOVA on RT confirmed that there were reliable main effects of block ($F(4,88)=27.7$, $p<0.001$) and stimulus type ($F(2,44)=33.0$, $p<0.001$), though the interaction of these factors was not significant. The significant effect of stimulus type followed from reliably slower responses in the cohort condition than in rhyme or unrelated conditions (which did not differ from each other).

Accuracy and RT are less differentiated among the trial types in the uncorrelated condition. For accuracy, the main effect of stimulus type was not significant ($F(3,69) < 1$), but there was a significant effect of block ($F(4,92)=62.4$, $p<0.001$) and a significant interaction of block and stimulus type ($F(12,276)=1.8$, $p<0.05$). Post-hoc tests confirmed that the interaction followed from reliably lower accuracy in v1p0 in blocks 1 and 2 and for v0p1 in block 4. For RT, block was significant ($F(4,92)=6.1$, $p<0.001$), as was the main effect of stimulus type ($F(3,69)=4.1$, $p<0.05$). Post-hoc tests confirmed that the latter effect was due to reliably faster responses in v0p1 than in v2p0 and v1p0, indicating that visual overlap inhibited learning more than phonological overlap.

Note that the same similarity relationships cannot apply in correlated and fully uncorrelated conditions. To achieve full absence of correlation, items can overlap in cap, stem, first syllable, or second syllable, but never in two of these features. A trend worth noting is the step-like function present for visual overlap trials, as the accuracy was lowest in the first two blocks and rose sharply to be among the highest for the last two blocks.

*Table 2: Correlations between experimental and assessment tasks. RST=Reading Span Task, SWE=TOWRE sight word efficiency, PDE=TOWRE pseudoword decoding efficiency; Faces, Objects, and Words = recognition memory in those domains (using d' as a measure of sensitivity). Correlations reliable at p<0.01 are bold with "**", p<0.05 are bold with "*", and with p<0.10 are bold with "+".*

		RST	*SWE*	*PDE*	*Faces*	*Objects*
Overall	*SWE*	**0.47****				
	PDE	**0.42****	**0.50****			
	Faces	0.23	0.13	**0.43****		
	Objects	0.17	0.11	-0.01	**-0.35***	
	Words	**0.27+**	**0.34***	**0.38****	0.13	**0.63****
Uncorrelated	*SWE*	**0.56***				
	PDE	**0.49***	**0.49***			
	Faces	0.33	0.22	0.33		
	Objects	0.01	0.14	-0.05	**0.38+**	
	Words	0.26	**0.41***	**0.41***	**0.48***	**0.54****
Correlated	*SWE*	0.32				
	PDE	0.24	**0.54***			
	Faces	**0.42+**	-0.04	0.06		
	Objects	0.34	0.10	0.02	**0.52***	
	Words	0.28	0.27	0.38	**0.56****	**0.70****

Individual differences

Table 2 presents correlations among the assessment scores. Results are first presented collapsed across correlation conditions, then by uncorrelated and finally by correlated condition. Performance on the recognition tasks was quantified as *d'* (sensitivity).

The first thing to note is that correlations are generally weaker for the correlated condition. Unsurprisingly, the two TOWRE subtests, SWE and PDE, correlate with one another. Also unsurprisingly, there are strong mutual correlations among the memory measures. Interestingly, better face recognition is correlated with better RST performance, a measure of verbal working memory.

In the uncorrelated condition, many more relationships emerge. Language measures are more strongly mutually related (including RST with the TOWRE subtests), as are memory measures. Additionally, we found significant correlation between spoken word recognition memory performance and the TOWRE subtests.

There is no reason to expect the correlated and uncorrelated groups to differ on relationships other than those with artificial lexicon measures. While similar associative trends are apparent for most pairs of measures for both groups, there are a few differences involving the recognition memory measures. It is possible that these differences are the result of greater fatigue in the more challenging uncorrelated condition by the time participants completed the recognition memory tests.

When we seek to increase power by collapsing across conditions, a few additional details emerge. Language measures correlate strongly with one another. Face and object recognition correlate with one another, as do object and spoken word recognition. However, face and word recognition do not correlate significantly with one another. Curiously, face and object recognition correlate negatively. Face recognition correlates strongly with the TOWRE PDE subtest, while spoken word recognition memory correlates with the TOWRE subtests (and is approaching significance with RST).

Recall that our goal in assessing individual differences in language and memory was to test predictions that follow from Perfetti and Hart's (2002) Lexical Quality Hypothesis. If variation in reading ability cascades from differences in the strength and organization of lexical representations, then performance in a lexical learning task should relate to basic assessments of reading-related skills.

To test this, we used multiple regressions to explore potential relationships between task and assessment performance (specifically, average response time in the final experimental block, and average accuracy over all experimental blocks), with all predictors entered simultaneously. Table 3 summarizes the results of the separate analyses we conducted on RT and accuracy for the correlated and uncorrelated conditions. None of the models we tested was significant (though the model for correlated RT was marginally reliable).

For the purposes of exploring this data and its

Table 3: results of multiple regressions with all predictors entered simultaneously. Reliable and marginally reliable predictors are indicated by bold font.

Accuracy
Correlated Condition
$F(6,14) = 1.101, p = 0.409$

	Std. beta	*t*	*Sig.*	*Partial correlation*
RST	.463	1.716	**.108**	.417
SWE	.280	1.002	.333	.259
PDE	-.220	-.750	.466	-.196
Objects	-.476	-1.387	.187	-.348
Words	.234	.601	.557	.159
Faces	.003	.009	.993	.002

Response time
Correlated Condition
$F(6,14) = 2.178, p = 0.108$

	Std. beta	*t*	*Sig.*	*Partial correlation*
RST	-.200	-.849	.410	-.221
SWE	-.132	-.543	.596	-.144
PDE	-.025	-.097	.924	-.026
Objects	.763	2.548	**.023**	.563
Words	-.204	-.601	.557	-.159
Faces	-.541	-2.036	**.061**	-.478

Accuracy
Uncorrelated Condition
$F(6, 18) = 2.079, p = 0.107$

	Std. beta	*t*	*Sig.*	*Partial correlation*
RST	.246	1.039	.312	.238
SWE	-.365	-1.521	.146	-.337
PDE	.482	1.991	**.062**	.425
Objects	-.143	-.603	.554	-.141
Words	.141	.539	.596	.126
Faces	.144	.651	.523	.152

Response time
Uncorrelated Condition
$F(6, 18) = 0.847, p = 0.551$

	Std. beta	*t*	*Sig.*	*Partial correlation*
RST	.007	.027	.979	.006
SWE	.277	1.004	.329	.230
PDE	.157	.565	.579	.132
Objects	-.264	-.964	.348	-.222
Words	.056	.187	.854	.044
Faces	.073	.288	.776	.068

implications for the Lexical Quality Hypothesis, we examined which predictors contributed reliably (or marginally reliably) within each model. Surprisingly, language assessments were not strongly related to artificial lexicon learning, with marginal contributions from RST for correlated accuracy and pseudo-word decoding efficiency for uncorrelated accuracy. In the only model that approached significance (correlated RT), the strongest predictors were recognition memory for objects and faces. A somewhat puzzling aspect of this result is that the correlation between objects and RT was positive; as sensitivity increased in object recognition, RT lengthened. The expected relationship (a negative correlation) held for face performance.

Discussion

In this study, we assessed participants' abilities on a number of visual and language-based tasks, and examined how they related to performance in an artificial lexical learning task. Our goal was to discover to what degree performance on standard language assessments and other nonlinguistic tasks was related to performance in our learning experiment, which stressed phonological-semantic (in this case, visual) relationships. Our starting point was the Lexical Quality Hypothesis (Perfetti & Hart, 2002), which posits that variation in reading ability follows from variation in the strength and detail of lexical representations, and consequent differences in efficacy of lexical access. Unsurprisingly, performance in the experiment was better when the nature of the relationships was relatively transparent (when visual and phonological features were correlated). The absence of correlation in the uncorrelated condition resulted in slower learning and slower access. These results are in line with the Lexical Quality Hypothesis, as the mutually reinforcing correlations minimized competition and therefore potentially sped learning and facilitated access.

But what predicts individual variation in learning within these conditions? Our individual differences analyses provide some intriguing possibilities (though they must be considered with caution, given the weak overall regression results). Language assessments were less correlated with artificial lexical learning than one might have expected, with just two measures showing marginal predictions of accuracy. On the other hand, object and face recognition performance were modest predictors of RT in the correlated condition. The correlation with face recognition is intriguing: intuition would suggest that, of the assessments we used, face recognition would be the least related to the experimental task. What connection might there be between these seemingly distant aspects of memory and learning?

Consider the fact that the correlated condition is where highest efficiency is achieved in the artificial lexicon task. Perhaps object and face recognition predict RT in the correlated condition because they are indexing domain-general memory access efficiency. If this were the case, though, we might also expect RST to predict correlated RT. It may be instead that object and face recognition predict correlated RT because all three index efficiency of configural learning and processing.

Our next steps will include testing these speculations about the relationship of face recognition and lexical learning using neuroimaging. We will also extend the current approach with artificial lexicons where visual referents are orthographic rather than pictorial, in order to test whether these results generalize to orthography. Similar results might suggest that one aspect of linguistic learning -- linking material across modalities -- may be a source of specific weakness in some learners.

Acknowledgments

Support: NIH Grant HD001994 to Haskins Laboratories.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247-264.

Dahan, D., & Tanenhaus, M.K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review*, *12*, 453-459.

Harm, M. W., & Seidenberg, M. S. (1999). Reading acquisition, phonology, and dyslexia: Insights from a connectionist model. *Psych. Review, 106*, 491-528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes. *Psych. Review, 111*, 662-720.

Magnuson, J.S., Tanenhaus, M. K., Aslin, R.N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General, 132*, 202-227.

Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory & Cognition, 37*, 1026-1039.

Perfetti, C.A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of functional literacy* (Vol. 11). Amsterdam/ Philadelphia: John Benjamins Publishing.

Richardson, D.C., Dale, R., & Spivey, M.J. (in press). *Eye movements in language and cognition: A brief introduction.*

Tanenhaus, M. K., & Spivey-Knowlton, M. J. (1996). Eye-tracking. *Language & Cognitive Processes, 11*, 583-588.

Tanenhaus, M. K., Spivey Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632-1634.

Torgesen, J. T., Wagner, R. K., & Rashotte, C. A. (1999). Test of word reading efficiency (TOWRE). Austin, TX: PRO-ED.