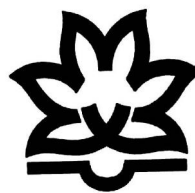


Content list available at www.urmia.ac.ir/ijltr

*Iranian Journal
of
Language Teaching Research*



Urmia University

Reading and vocabulary acquisition: supporting evidence and some objections

Stephen Krashen^{a,*}

^a *University of Southern California, US*

ABSTRACT

I continue here the long-standing discussion on the familiar topic of whether subconscious language acquisition is more powerful than conscious language learning, with a focus on vocabulary, adding recent studies as well as older ones I missed in previous publications on this topic (e.g. Krashen, 2004).

Keywords: language acquisition; language learning; vocabulary acquisition; reading

© Urmia University Press

ARTICLE SUMMARY

Received: 7 Dec. 2012

Revised version received: 11 Dec. 2012

Accepted: 18 Dec. 2012

Available online: 25 Dec. 2012

* Corresponding author: University of Southern California, US
Email address: skrashen@yahoo.com

Before starting in, I present what I find the most disturbing argument presented in support of conscious learning and direct instruction.

The Brain versus Society's Expectations

A common practice is to describe some aspect of language that first or second language acquirers have not yet acquired or fully acquired, but one that people think the students must know to meet society's demands. It is then declared that the items or rules must be taught directly.

Evans and Green (2007), for example, present evidence that students and faculty in higher education in Hong Kong agree that students have deficits in understanding academic English vocabulary. They then recommend classes in English for Academic Purposes as well as more testing in these areas (See also Hyland, 2000).

What is not considered is the possibility that these aspects of language are late-acquired and will be acquired if the student gets more comprehensible input. If it is an aspect of academic language, it will be acquired by reading.

The method of finding out if an item is late-acquired and can be acquired by the acquirer from comprehensible input is simple: Look at more advanced acquirers and see if they have acquired the item. If the results are variable, if some have acquired it and some have not, we can determine why by asking about their experiences, specifically how much reading they have done. This kind of investigation will not only inform us whether the item will be acquired, but will also give us some expectations as to what it takes.

This is never done. Instead, it is assumed that these items should be taught, an approach that has never worked.

THE CASE FOR READING

In this section, I briefly review the arguments supporting the view that vocabulary is acquired as a result of receiving comprehensible input, adding some recent research as well as older studies I missed. For academic vocabulary, this means reading.

Read and test studies

In these studies, subjects read a passage that contains words unfamiliar to them, are not focused on the new words while reading, and are given a surprise test on the words after reading the passage. In both first and second language studies, readers show clear evidence of increasing their knowledge of these unfamiliar words after having encountered them in print (reviewed in Krashen, 2004).

Read and test studies not only tell us that it is possible to gain vocabulary knowledge from reading, they also tell us that the process is gradual and that quite a bit of reading is required to build a sizable vocabulary. Nagy, Herman, and Anderson (1985) and Nagy, Anderson, and Herman (1987) used measures that were sensitive to whether subjects had acquired just part of the meaning of a target word. Nagy and colleagues concluded from their data that when an unfamiliar word was seen in print, 'a small but reliable increase of word knowledge' typically occurred (Nagy & Herman 1987, p. 26), but this increase was enough to account for vocabulary acquisition given enough reading, about a million words a year for fifth graders.

Second language research reaches a similar conclusion, based on findings showing that read and test experiences can produce transient results. Waring and Takaki (2003) did not use real words, but substituted nonsense spellings for real words in a text (e.g., 'yes' became yoot'). Subjects did well on tests given immediately after the reading, but scores dropped on a test given three months later. The low scores on the delayed test were not enough to account for vocabulary growth. One reason for this is the fact that subjects did not encounter any of the target words during the interval

between the treatment and the delayed tests, because the words were artificial and do not occur in normal English. Waring and Takaki conclude that vocabulary growth requires many encounters with unfamiliar words – in other words, a great deal of reading (see also Brown, Waring & Donkaewbua, 2008).

An objection: Medina, Snedeker, Trueswell and Gleitman, (2011) claim that vocabulary acquisition is ‘all or nothing,’ each word acquired all at once rather than gradually.

In their study, adults observed brief scenes of parent-child interaction. The target words were replaced by beeps or nonsense words, and subjects were asked to guess the meanings of the target words. It was first determined that in only 7% of the scenes was it obvious what the target word meant, that is, only 7% were ‘transparent.’

If subjects saw a transparent scene early, they often guessed the meaning correctly and did not change their mind with subsequent scenes. But if the first scene was not transparent, subjects did not discover the meaning of the missing word, did not get closer from seeing more nontransparent scenes, and, in fact, did not remember their guesses a few days later.

My interpretation is that this study did not probe real vocabulary acquisition, but probed conscious learning. The subjects were well-educated college students. During the study, it was clear that the subjects were aware that they were trying to guess the meanings of the missing words and were trying to remember them.

The missing words were not presented as part of highly interesting, comprehensible input; rather, subjects observed short, random scenes of parent-child interaction. Most of the input was, most likely, incoherent and boring.

Thus, Medina et al. really found that for educated adults, transparent input results in conscious learning and that non-transparent input or less transparent input does not contribute to conscious learning.

Sustained Silent Reading

In Sustained Silent Reading (SSR), students are given some time each day to do self-selected reading with no or minimal accountability. SSR students typically do better than comparisons in vocabulary development (Krashen, 1989, 2004), and this has been confirmed in a number of second language studies (Lee, 2007; Lee & Hsu, 2009; Rodrigo, Krashen & Gribbons, 2004).

Print exposure

A meta-analysis involving 99 studies strongly confirms previous results. Mol and Bus (2011) reported consistent and strong correlations between print exposure, as measured by author and title recognition tests, and aspects of literacy and language development. Title and author recognition tests ask respondents to indicate whether or not they are familiar with the titles of books or authors. A number of studies by Keith Stanovich and others have shown that scores on author and title recognition tests are related to the amount of reading done, as well as observed reading (reviewed in Krashen, 2004, pp. 11-12).

Table 1 summarizes Mol and Bus' results in terms of effect sizes (d). Clearly, the richness of the print environment is positively related to literacy competence: Those who recognize more titles and/or authors and thus, we assume, have read more and live in more print-rich environments, have larger vocabularies, read better, spell better, and have more phonological awareness.

Table 1: Results of Mol and Bus meta-analysis

measure	prek-k	Grades 1-12	college
oral vocab	.72 (12)	1.12 (18)	1.74 (18)
reading comprehension		.82 (21)	.97 (11_)
phonological processing		.44 (14)	.38 (5)
Spelling		.82 (7)	.93 (14)
basic reading skills	.6 (8)		

Parenthesis = the number of studies the effect size is based on.

College: combines studies involving undergraduate and graduate students

Mol and Bus used r as a measure of effect size; r converted to d by SK.

Mol and Bus also reported that 'less proficient readers' (including second language acquirers) profited from print exposure.

An objection: Lawrence (2009) studied low-SES (socio-economic status) fifth and sixth graders' summer reading, and concluded that 'independent reading is not sufficient to keep the word-learning faucet 'on' during the summer months' (p. 462).

Lawrence divided his subjects' reported reading into four categories: narrative (novels, poetry, biography), expository (information books, research reports, instructions, map, bus airline schedules), 'teen' (comic books, magazines, music lyrics), and computer (email, websites). He reported no relationship between the amount of each type of reported reading and gains in vocabulary during the summer. In fact, the amount of 'teen' reading done was negatively correlated with vocabulary growth.

It is clear from Lawrence's report that these students were not fanatic pleasure readers. Being low SES, they did not have a great deal of access to reading material. Nevertheless, Lawrence's analysis does not really take into account what they did read, because it dealt with only one genre at a time, not the total amount read. There

was nowhere near enough reading done in any of the genres individually to allow a correlation to emerge.

For example, Lawrence failed to find a correlation between reported 'narrative' reading and vocabulary growth. But subjects reported only an average of 13 sessions of reading narrative texts throughout the entire summer, about one reading session a week. Combining all reading, all genres and subgenres (from Lawrence's table 3), result in an average of 109.5 reading sessions over the summer. If Lawrence had examined the relationship between vocabulary growth and all reading, it is much more likely that a positive and significant correlation between reading and vocabulary growth would have emerged, if indeed it actually existed in reality.

Even if the resultant correlation were weak, the more logical path, based on the research, is to increase access to reading material, not rush to do direct instruction.

Size and Complexity

In previous publications, I discussed arguments that direct instruction cannot deal with the size and complexity of vocabulary learning: There are simply too many words to be acquired (Krashen, 2004; see Nation (2006) for an analysis of vocabulary demands of different kinds of input). Hyland (1996) presents an excellent example of the complexity of vocabulary. Hyland presents data showing that second year business students at a Hong Kong university have not fully acquired the subtleties of 'quite.'

Hyland presents a thorough discussion of the complexity of 'quite,' e.g. it is both a 'booster' (e.g. 'the results were quite phenomenal') and a 'hedge' or slight attenuation (e.g. 'he couldn't quite do it'), but after this simple generalization, things get 'fuzzy,' as Hyland points out. 'Quite' varies in meaning according to stress, e.g. 'I QUITE like the idea of walking' (but I'd prefer not), versus 'I quite LIKE the idea of walking' (and maybe I will), and whether it comes before or after the article, e.g. 'a quite

beautiful garden' versus 'quite a beautiful garden,' the former expressing 'greater commitment.' He also notes that pedagogical grammars as well as professional linguists differ in their rules for 'quite' and discusses the inadequacies in their presentations.

Hyland acknowledges that the 'pragmatic complexity' of 'quite' means that it cannot be taught in the usual way: '... the fact that linguists differ in their preferred accounts of its meanings and implications means that classroom activities based on textbook exercises or intuition-based grammars are unlikely to lead to a clear understanding. As a result, there are obvious advantages in turning to authentic examples' (p. 103).

But just when we think Hyland is about to embrace reading, he turns the other way: His solution: 'turn students into researchers' (p. 104), and have them examine the use of 'quite' in different corpora. In other words, study the use of 'quite' in order to arrive at generalizations. Students, Hyland suggests, can work in groups on such issues as the meaning of 'quite' with or without an article, 'reach a consensus ... and provide justifications for their decisions' (p. 104), and then cheerfully go on to compare the use of 'quite' in different genres, spoken and written texts, etc. Hyland even suggests that students write research papers on this topic.

There are several problems with this solution. First, it assumes students would be interested in becoming junior linguists. Second, these activities would not lead to acquisition, but to conscious learning. Third, even if conscious learning were worthwhile, the rules that students would come up with would probably not be any better than those in the texts and in the professional literature. Fourth, 'quite' is only one tiny aspect of English. Doing research on all complex aspects of language would take several lifetimes.

Again, there is no discussion of the possibility that students' acquisition of 'quite' will improve as they get more comprehensible input. This is easy to investigate: Did

those with better performance on Hyland's 'quite' test do more reading? Do more advanced English acquirers have higher 'quite' competence?

Read-alouds/storytelling and vocabulary growth

A number of studies confirm claims that read-alouds and story telling are a powerful source of vocabulary.

Studies have shown that storybooks are a rich source of vocabulary, far richer than textbooks (Hsieh, Wang & Lee, 2011), and that reading aloud stimulates the desire and ability to do self-selected reading (Brassell, 2003; Cho & Choi, 2008; Hsieh et al., 2011; Trelease, 2006; Wang & Lee, 2007).

In a series of studies of adult second language acquirers, Beniko Mason (e.g., Mason & Krashen, 2004) concluded that developing vocabulary knowledge from listening to stories is more efficient in terms of words acquired or learned per minute than vocabulary-building exercises.

Reach Out and Read: Reach Out and Read (henceforth ROR), an inexpensive program based on reading aloud to very small children, has produced encouraging results.

Reach Out and Read is simple: While in waiting rooms for well-child pediatrician's appointments, hospital staff show parents reading activities they can do with their children, with a focus on reading aloud to the child, and discuss the importance of reading, which the physician does as well. The families receive free books at each doctor visit. ROR is typically aimed at lower-income groups.

I present the results of a typical ROR study here; others are presented in Krashen (2011). In Mendelsohn, Mogiler, Dreyer, Forman, Weinstein, Broderick, Cheng K., Magloire, Moore and Napier (2001), children were about four years of age and had had three years of experience with ROR. Two kinds of vocabulary tests were used:

Expressive and Receptive One-Word Picture Vocabulary tests. As was the case in other ROR studies, ROR children did better on the vocabulary measures and the ROR advantage was larger on receptive than expressive vocabulary tests.

Table two: Results of Mendelsohn et. al.

	Expressive	Receptive
comparison (n = 49)	80.9	85.2
ROR (n = 73)	85.2	93.7
effect size	.29	.57
national norm	100	100
Gap	19.9	14.8
% gap closed	4.3/19.9=22%	8.5/14.8=57%

Means adjusted for differences between the groups, e.g. mother's education, language spoken in the home, homelessness, preschool attendance, child's age. Test scores standardized for age (100 = 50th percentile).

Effect size calculated from means and standard deviation (sd=15).

Because all subjects were from low socio-economic families, it is not surprising that the children scored below the national median (100). The ROR children, however, closed from about one-fourth to one-half of the gap on the receptive test.

A modest treatment: The entire treatment consisted of a few well-child pediatrician visits, providing some information about reading aloud to children, and providing a small number of books. Over a three-year span, subjects in Mendelsohn et al. had an average of only three well-child appointments in which their doctors discussed books and they received an average of four books.

The importance of read-alouds: Studies of ROR consistently show that ROR children are read to more than comparison children. In Mendelsohn et al., Reach out and Read children were read to an average of 4.3 times per week, while comparisons were read to only 2.8 times per week, a statistically significant difference.

The crucial role of read-alouds was confirmed by High, LaGasse, Becker, Ahlgren, and Gardner (2000): Frequency (days per week) of read-alouds was a strong predictor of scores on both vocabulary measures, controlling for demographic variables,

including parental language proficiency. In fact, when frequency of read-alouds was considered, High et al. reported that participation in ROR had no additional impact on vocabulary test performance.

An objection: Marulis and Neuman (2010) performed a meta-analysis of studies on vocabulary instruction among pre-k and kindergarten children, and reported that the average effect size for studies using explicit training (direct instruction) was larger (Hedge's $g = 1.11$, 15 comparisons) than for studies using implicit training (without deliberate teaching) ($g = .62$, 25 comparisons). Marulis and Neuman conclude that explicit training is more effective than letting children pick up the meanings of words from context, without focusing on the words.

This is not evidence against read-alouds. The effect size of .62 found for implicit methods means that the implicit approach was indeed effective. Children who were exposed to new words in context, without instruction, clearly did quite a bit better in vocabulary development than control subjects. It is highly likely that most of the 'implicit' methods were read-alouds and/or storytelling. In addition, very few storybook studies showed near-zero or negative effect sizes.

Four reasons why storytelling and read-alouds are preferable to direct (explicit) instruction

ONE: An explicit vocabulary treatment focuses only on vocabulary. Read-alouds and storytelling clearly stimulate substantial vocabulary growth, but have great value other than vocabulary acquisition in terms of language, cognitive and psychological development (e.g., Trelease, 2006). Moreover, explicit vocabulary teaching focuses only on a restricted list of vocabulary words: Stories also contain a rich supply of vocabulary as well as grammar and cultural information.

TWO: Explicit instruction might have only a short-term effect. In Marulis and Neuman, in all studies post-tests were given immediately after the treatment. All data

from delayed posttests were excluded. Thus, the superiority of explicit instruction might not be long-lasting.

THREE: Read-alouds and storytelling require only a modest amount of preparation. Explicit vocabulary lessons must be planned in great detail.

FOUR: Read-alouds and storytelling are pleasant for both children and readers. They are activities parents/caretakers and children engage in willingly and enthusiastically for years. It is doubtful that there would be such enthusiasm for explicit vocabulary teaching.

AN ADDITIONAL COMMENT

I conclude with a brief discussion of what some of the problems are in doing research in this field that, to my knowledge, have not been discussed in print before.

Even this short review reports on quite a bit of research in diverse areas: Vocabulary studies have been done in first and second language development, and for this reason are hard to find. I had missed Hyland's work that I discussed above, and was unaware of the Reach Out and Read studies, for example. We are not dealing with just one or two central journals.

Another problem is that getting access to information from so many sources can get expensive for those without a free university service. In addition, the price of professional books is now out of control, with edited volumes often selling for over \$100 US.

The price of getting information makes it hard not only for consumers of research, but it also nearly insures that scholars' efforts will have nearly zero impact. Publishing in an expensive volume condemns work to obscurity. It helps that authors put their papers on their websites, but first readers have to find out that the articles exist.

A third problem: Most papers are much too long. Sometimes we have to write long papers to treat a topic adequately, but much of the time, it's unnecessary: Papers often contain long introductions more suitable for doctoral dissertations, apparently intended to provide evidence that the author is well-read, and long conclusions, with a repetition of the findings and the author's detailed and lengthy speculations about what the results might mean for theory and application.

Readers of professional journals don't need this. Introductions need to only give enough to alert the reader to what the article is about, and provide a few citations in case the reader needs more information. In addition, a brief indication of implications is generally more than enough for experienced readers. Also, if the results section is clear, no repetition of the findings is necessary in the conclusion.

Unnecessarily long papers take too much time to write and of course much too much time to read. They also take up space in journals, which hurts the dissemination of knowledge, because less genuine information is available, and makes it much harder to publish, a serious problem especially for junior scholars.

Language education has clearly taken its tradition from the humanities, which favors dissertation-style prose, rather than the sciences, where papers are usually much shorter.

It is probably no coincidence that citation rates in the sciences are much higher than in language and linguistics: Hamilton (1991) reported that about 91% of papers published in atomic, molecular and chemical physics, and 86% in virology had been cited at least once. In language and linguistics, only 20% had been cited and in American literature, less than 1%.

Watson and Crick's Nobel Prize winning paper on the double helix (Watson & Crick, 1953) was only one page long.

Some recommendations

More free, open-access journals, such as *Reading in a Foreign Language*, the *International Journal of Foreign Language Teaching*, the *Iranian Journal of Language Teaching Research*, and *Language and Language Teaching*.

When possible, shorter papers. ‘When we ask the time, we don't want to know how watches are constructed.’ Georg Christoph Lichtenberg (1742-1799)

References

- Blok, H. (1999). Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning* 49(2), 343-371.
- Brassell, D. (2003). Sixteen books went home tonight: Fifteen were introduced by the teacher. *The California Reader* 36(3), 33-39.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136-163.
- Bus, A., Van Ijzendoorn, M., & Pellegrini, A. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65, 1-21.
- Cho, K.S., & Choi, D.S. (2008). Are read-alouds and free reading ‘natural partners’? *Knowledge Quest*, 36(5), 69–73.
- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6, 3-17.

- Hyland, K. (1996). 'I don't quite follow': Making sense of a modifier. *Language Awareness* 5(2), 91-109.
- Hyland, K. (2000). Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness* 9(4), 179-197.
- High P., LaGasse L., Becker S., Ahlgren, I., & Gardner, A. (2000). Literacy promotion in primary care pediatrics: Can we make a difference? *Pediatrics*, 104, 927-934.
- Hamilton, D. (1991). Research papers: Who's uninvited now? *Science*, 251, 25.
- Hsieh, M.Y., Wang, F.Y., & Lee, S.Y. (2011). A corpus-based analysis comparing vocabulary input from storybooks and textbooks. *International Journal of Foreign Language Teaching*, 6(1), 25-33.
- Krashen, S. (2004). *The power of reading*. Portsmouth: Heinemann and Westport: Libraries Unlimited.
- Krashen, S. (2011). Reach out and read (aloud). *Language Magazine*, 10(12), 17-19.
- Lawrence, J. (2009). Summer reading: Predicting adolescent word learning from aptitude, time spent reading, and text type. *Reading Psychology*, 30, 445-465.
- Lee, S.Y. (2007). Revelations from three consecutive studies on extensive reading. *RELC Journal*, 38(2), 150-170.
- Lee, S.Y., & Hsu, Y.Y. (2009). A three-year longitudinal study of in-class sustained silent reading with Taiwanese vocabulary college students. *Indonesian Journal of English Language Teaching*, 5(1), 15-29.

- Marulis, L., & Neuman, S. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research* 30(3), 300-315.
- Mason, B., & Krashen, S. (2004). Is form-focused vocabulary instruction worth while? *RELC Journal* 35(2), 179-185.
- Medina, T., Snedeker, J., Trueswell, J., & Gleitman, L. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Science*. 108(22), 9014-9019.
- Mendelsohn, A., Mogiler, L., Dreyer, B., Forman, J., Weinstein, S., Broderick, M., Cheng, K., Magloire, T., Moore, T., & Napier, C. (2001). The impact of a clinic-based literacy intervention on language development in inner-city preschool children. *Pediatrics*, 107(1), 130–134.
- Mol, S., & Bus, A. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267-296.
- Nagy, W., Anderson, A., & Herman, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24, 237-270.
- Nagy, W., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown, M. Curtiss, & M. Hillsdale (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). NJ: Erlbaum.
- Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly* 17, 233-255.
- Nation, P. (2006). How large a vocabulary is needed for listening and reading? *Canadian Modern Language Review*, 63(1), 59-82.

Rodrigo, V., Krashen, S., & Gribbons, B. (2004). The effectiveness of two comprehensible-input approaches to foreign language instruction at the intermediate level. *System*, 32(1), 53-60.

Trelease, J. (2006). *The read-aloud handbook*. New York: Penguin. Sixth Edition.

Wang, F. Y., & Lee, S. Y. (2007). Storytelling is the bridge. *International Journal of Foreign Language Teaching*, 3(2), 30–35.

Watson, J., & Crick, F. (1953). A structure for Deoxyribose nucleic acid. *Nature*, 171, 737-738.

Stephen Krashen is professor emeritus at the University of Southern California. He is a linguist, educational researcher, and activist. He has published more than 350 papers and books, contributing to the fields of second-language acquisition, bilingual education, and reading. He is credited with introducing various influential concepts and terms in the study of second-language acquisition, including the acquisition-learning hypothesis, the input hypothesis, the monitor hypothesis, the affective filter, and the natural order hypothesis.