

Text signals influence second language expository text comprehension: knowledge structure analysis

Kyung Kim¹ · Roy B. Clariana¹

Published online: 12 October 2016

© Association for Educational Communications and Technology 2016

Abstract This quasi-experimental investigation describes the influence of text signals on second language expository science text comprehension. In two course sections, mixed proficiency Korean English language learners ($n = 88$) read one of two print-based English expository text passage versions. Participants in one section ($n = 44$) were given a version with interesting but *non-important subtopics* (NIS) underlined, while participants in the other section ($n = 44$) were given an alternate version with the same number of *substantively important subtopics* (SIS) underlined. Participants read the text passage and created a visual map of the text, and then completed a comprehension posttest that measured global inferences (all in English). Analysis of variance of the comprehension posttest data revealed significant differences for the two main factors, proficiency level (Low or High) and text signal condition (NIS or SIS), and for the interaction of proficiency and text signal condition; with Cohen effect sizes, $d = \text{Low-NIS} (.16) < \text{High-NIS} (.56) < \text{Low-SIS} (.74) < \text{High-SIS} (1.37)$. Descriptively, both headings and important subtopic terms predominated in the SIS maps, but only non-important subtopic terms and not headings predominated in the NIS maps. Further, the visual map forms, as measured by vector pattern matching and by graph centrality, were also substantially different. Compared to the NIS maps, on average the SIS map forms were more relational and more like the expert's map, while the NIS map forms were more linear, and showed a primacy effect. These results suggest that coherent text signals in these print-based readings strongly influenced bilinguals' science expository text comprehension.

Keywords Text signals · Expository text · Second language reading · Knowledge structure analysis · Instructional design

✉ Kyung Kim
kxk997@psu.edu

¹ Learning, Design, and Technology, The Pennsylvania State University, University Park, PA, USA

Introduction

The U.S. Census Bureau's American Community Survey (Ryan 2013) reported that the number of people who speak a language other than English at home is 61.8 million, up 2.2 million since 2010. The largest increases from 2010 to 2013 were for speakers of Spanish, Chinese, and Arabic. Thus currently, one in five U.S. residents now speaks a foreign language at home (e.g., as their first language). What instructional design advice is there for these students in our classrooms who must learn in English? Regarding text design for reading comprehension (Jonassen 1985), instructional designers and authors should use text signals to more clearly indicate the domain-specific knowledge structure (Lemarié et al. 2012), but the influence of such signals on second language readers is not well understood. Thus, it is important for instructional designers to know whether and how the text signals influence individual reader's understanding of the text in any way. This investigation positions the likely effects of text signals within second language expository text reading.

Text signals in reading

Various text signals have been investigated in reading research, such as titles, headings, logical connectives (e.g., however, as a result, in fact, etc.), and typographical cues (e.g., underline, content spacing, bold facing, etc.). Native language (L1) reading studies have demonstrated that text signals facilitate processing and comprehension of a text, especially of an expository text, by directing readers' attention to the important word/topic/phrase that the author intended [e.g., Clariana et al. (2015), Lemarié et al. (2012), Lorch et al. (2011), Meyer et al. (2012)], and thus establishing a coherent mental text representation [the "situation model" of van Dijk and Kintsch (1983)] that is consistent with author's situation model—the ultimate goal of expository texts.

That is, text signals support readers' ability to identify key textual concepts and organize these ideas in a coherent manner, especially for readers who have trouble in understanding the text, such as readers in a second/foreign language (L2). However, the bulk of L2 reading studies have investigated the impact of text signals on learners' acquisition of lexical elements only [i.e., parts of speech, imperatives, subjunctives, participle agreement, passive voice, and many others; see LaBrozzi (2016), Wong (2003)] but have not considered comprehension (i.e., semantic level). In this investigation, thus, we explored the likely influence of text signals on comprehension of L2 expository text.

Text signals and L1 reading

A number of reading studies (note: with monolinguals, L1 only) have provided ample evidence that generally the use of text signals has a positive influence on readers' comprehension [see Lemarié et al. (2012) for a review]. Certain moderating findings have emerged from these previous L1 studies that may or may not apply in L2 settings: (a) signals have the greatest effects when the texts match the proficiency level of the readers, that is, the texts are neither too difficult nor too easy [e.g., Spyridakis and Standal (1987)]; (b) signals aid readers best when the texts are unfamiliar technical or scientific topics, particularly texts written with the goal of conveying information to the readers who are unfamiliar with the content, i.e., expository texts [e.g., Degand and Sanders (2002)]; and (c) each signal would best work alone rather than in combination because each signal

device serves a distinct information function with distinct implications for text processing [e.g., Lorch et al. (2011)]. The present investigation continues the previous lines of L1 reading research and explores the effect of text signals on L2 text comprehension by limiting the scope to a specific type of signal, *underline*, with an expository scientific text that is a relatively difficult text on an unfamiliar topic.

L1 reading studies have established the effectiveness of text underlining on expository text comprehension; when text information is underlined (i.e., isolated against a homogenous background), readers will recall that signaled information better than non-underlined information, supporting the von Restorff effect, and then perform well on exams which test recall of that specific information [e.g., Cashen and Leicht (1970), Crouse and Idstein (1972), Fowler and Barker (1974), Hartley et al. (1980), Nist and Hogrebe (1987)]. But there has been little research on the effect of text underlining on expository text comprehension in L2 context.

Text signals and L2 reading

In an L2 context, most studies have investigated the impact of the text signals on L2 learners' acquisition of form only (i.e., grammar learning), typically using typographical signals (e.g., underline, italicization, bolding, shading, etc.) to enhance the saliency of the form [see Lee (2007) for review], and the findings are inconclusive; some studies reported positive impacts of typographical signals on acquisition of targeted forms [e.g., Izumi (2002), Shook (1994)] while the other studies did not [e.g., Leow (2001), Overstreet (2002)]. Wong (2003) properly argued that "...the role of text signals in L2 cannot be complete without information about how comprehension is affected (or not affected) as learners' attention is directed at signal" (p. 21). The present investigation is designed to address this previous limitation of the research base, and to provide insight into how text signals might contribute to L2 expository text comprehension.

However, it is important to note that L2 proficiency significantly contributes to comprehension of L2 text, as evidenced by the bulk of L2 reading studies [e.g., Fecteau (1999), Lee and Schallert (1997)], so we assume that the effects of text signals will differ based on readers' L2 proficiency level. For example, low proficient L2 readers most likely will rely more on the text signals, in this case underlining, because their L2 language background is less able to grasp the topic structure of the text, while high proficient L2 readers may or may not use the signals depending on their reading goal, content familiarity, and other factors. Thus, this study explores (1) how text signals impact L2 readers' expository text comprehension, and (2) whether the supposed effects differ by their L2 proficiency level.

Knowledge structure and reading comprehension

Expository texts (e.g., scientific text) intend to describe the relationships between units of information mainly as sets of propositions locally, but also globally at the paragraph and section levels (Meyer et al. 1980). The primary communicative purpose of expository text is to 'inform' so that the readers learn something. Thus, the ultimate goal of reading expository scientific texts is to arrive at an appropriate understanding of the underlying domain-normative knowledge structure intended by the author/content expert. This knowledge structure (KS) is an important aspect of domain knowledge (Clariana 2010; Jonassen et al. 1993) that has been shown to relate to text comprehension in both L1 and L2 [e.g., Clariana et al. (2014), Kim and Clariana (2015), Clariana et al. (2015), Barry and Lazarte (2000), Meyer et al. (2012)]. Cognitive psychologists consider KS as a person-

specific factor, where a reader constructs an individualized KS of the text, a situation model that integrates new understanding from the text into the reader's prior knowledge base in memory (Kintsch 1988). The situation model is the reader's *interpretation* of the text, the personal mental model of what the text is about (Perfetti 1989). Our view is that the author's high-dimensional KS (the author's situation model) must be encapsulated into a lower-dimensional sequential text form, and then the reader reconstitutes the text information back into their own higher dimensional KS of that text (the reader's situation model). In this perspective, the reader's situation models may or may not match the author's situation model reflected in the text, but the reader's ability to form a coherent situation model that the author intended is an indicator of successful STEM-content reading comprehension (Fesel et al. 2015; Kintsch 1988; Zwaan and Radvansky 1998).

Then, how can we effectively capture the reader's situation model as KS in this cognitive perspective of reading? Production tasks of words and sentences are notoriously difficult for bilingual readers, so it is important to have a cognitively easy KS elicitation task that is comparable for bilinguals at different proficiency levels of L2 (van Hell and Kroll 2013). Recent cognitive studies emphasize KS as associative networks of concepts that contain weighted connections (much like a mental lexicon that contains associations between words), and reading allows the strengthening of the connections as well as the enrichment of concepts to occur in the network. Thus, methods that can capture network properties are most effective for describing the KS (Zareva and Wolter 2012).

Asking readers to make a visual is one way. For example, when given this *intentionally ambiguous* sentence adapted from Zwaan and Radvansky (1998), "A turtle rested beside a floating log, and a fish swam beneath it" (note: there is a 50–50 chance of guessing the author's intention), if a reader is then asked to use the terms 'turtle, log, and fish' to make a meaningful visual of the sentence, they may place the fish under the log or under the turtle; the visual will represent their individual *situation model* of this text; that is, how they have understood the text. In well written unambiguous expository texts, the text structure and the reader's situation model are more likely to be the same. Nevertheless, if the reader truly represents their understanding of the text in their visual, by definition, the visual is their individual situation model.

Keeping this in mind, this present investigation employed a very easy non-hierarchical visual mapping approach as a method to provide snapshots of situation model KS of the L2 text; for example, participants simply "sort" terms by moving related terms closer together and unrelated terms further apart, and then "link" the terms if they are strongly related to show direct relationships over-and-above proximity. Such visual maps in the form of node-link-node assemblies are a widely used paradigm in cognitive and educational psychology [e.g., Lambiotte et al. (1989)], brain research [e.g., van Hell and Kroll (2013)], and psycholinguistics [e.g., Zareva (2007)]. Recent studies have demonstrated that this spatial representation approach is a sensitive method for assessing the effects of reading in both L1 and L2 [e.g., Kim and Clariana (2015), Fesel et al. (2015)].

Purpose

Although L1 reading studies have established that text signals facilitate science expository text comprehension, the effects of text signals on comprehension have rarely been the focus of L2 reading studies. Thus, this current investigation explores how text signals influence L2 expository text comprehension and whether this differs by L2 proficiency level, and also considers KS complexity as a factor in comprehension because if readers comprehend the expository text as an author/expert intended, then the author's KS would

be reflected in the readers' KS. Thus, the term "comprehension" is operationally defined in terms of the correspondence (or lack thereof) of KS between a reader and the expert; for example, looking forward to the results of this study, "correct" text signal readers have better KS (more like the expert KS) and this better KS engenders better comprehension posttest performance, compared to "wrong" text signal readers.

Spector et al. (2015), the editors of this journal, have called for replication studies to support scientific rigor. This investigation replicates a recent investigation in this journal by Clariana et al. (2015) using the same treatment conditions (e.g., signaling *non-important subtopics* or *substantively important subtopics*) and uses the same descriptive analysis of students' post-reading maps, but with substantially different participants (e.g., Korean second language setting), different lesson content, and includes a well validated comprehension posttest measure (the TOEFL reading passage and test, *The Cave of Lascaux*, used with permission from the Educational Testing Service [ETS]).

Methods

Participants

This investigation was conducted with both high proficiency English Language Learners (ELLs) and low proficiency ELLs at a large Korean public university. There were two class sections heterogeneous by English proficiency, including 96 first-year students whose L1 was Korean. Students who had an official Test of English as Foreign Language [TOEFL, a valid and reliable measure of English proficiency, Laborda (2009)], were selected for this study, resulting in a sample size of 88 (section A $n = 44$ and section B $n = 44$). In both sections, there were $n = 23$ low proficient and $n = 21$ high proficient students (see Table 1). The average TOEFL iBT (internet-based test) score for Section A was $M = 42.88$, $SD = 2.42$, and for Section B was $M = 44.17$, $SD = 3.29$, the TOEFL iBT means for the two sections were not significantly different, $t = 1.549$, $p = 0.128$, confirming that entry English proficiency of the participants of the two classes was not significantly different. Regarding 'reading' score, the average TOEFL reading score for the high proficient participants was 26 out of 30 (ranged 22–30), which is regarded as *high* proficiency in English reading by ETS, while the average TOEFL reading scores for the low proficient group was 11 (ranged 5–14) that is categorized as *low* proficiency in English reading.

A web-based Language History Questionnaire [LHQ 2.0; Li et al. (2014)] was used to filter participants' proficiency level, this LHQ is popularly used in L2 studies for assessing the linguistic background of bilinguals and for generating self-reported proficiency. According to the LHQ, the high TOEFL score group revealed that thirty two of these individuals had previous experience studying English abroad for at least 6 years with

Table 1 The number of participants by condition and proficiency

Condition	Proficiency		
	Low	High	Mean
NIS (section A)	23	21	44
SIS (section B)	23	21	44
Mean	46	42	

different majors. Another ten high TOEFL score participants were graduate students who were majoring in English in the Korean university and who had stayed in the English-speaking countries such as the U.S., England, or Australia for at least 2 years for academic purpose. The LHQ for the low TOEFL score group indicated that all of the participants in this group did not have any intermediate or higher level English classes, nor did they have any previous experiences in English-speaking countries. All participants were briefed on the tasks involved and the purpose of this investigation and were requested to participate, and all agreed. They received course credits for the participation.

Materials

The materials were the ETS reading passage and associated multiple-choice comprehension posttest, *The Cave of Lascaux* (used with the permission of the ETS), a relatively difficult text on a generally unfamiliar scientific topic. This text had four paragraphs with headings, was 35 sentences long with 707 words, and a Flesch grade level readability/complexity score of about 13; thus it would be difficult for these L2 readers. For this study, following Clariana et al. (2015), we enhanced the passage with two different sets of underline signals: (1) the *non-important subtopic* signal version (NIS) has seven underlined terms to signal the interesting but *non-important subtopics* and (2) an alternate *substantively important subtopic* signal version (SIS) also with seven underlined terms to signal the essential important subtopic structure. Both versions of the text passage had four paragraphs with the same four prominent headings. These important/non-important subtopic terms were selected by three content experts (see Appendix for the text passage with signals).

Besides these 18 signaled terms mentioned above (i.e., 4 headings and 7 + 7 signaled terms/phrases), the open-ended maps created by participants included many other unsignaled terms. Following the approach used by Clariana et al. (2015), the frequencies of all of the unsignaled terms used by participants in their maps were calculated and the most frequently used eight terms were additionally included in this analysis (see these terms in the Appendix).

Procedure

Participants in section A ($n = 44$; 23 of low proficient, 21 of high proficient) received the NIS text version while those in section B ($n = 44$; 23 of low proficient, 21 of high proficient) received the SIS text version. First, participants completed a training lesson on how to draw a visual map (about 10 min); they were instructed to use any number of terms as they want in their maps (i.e., open-ended mapping). After the mapping lesson, they were asked to read the paper-based 707-word English TOEFL text passage, *The Cave of Lascaux* (either NIS or SIS version), and then create an English visual map of the text they read on the same paper handout (see for example Fig. 1). Immediately, they completed a multiple-choice posttest (from the ETS) that consists of 9 comprehension-level items to measure 'global' inferences of the text. The comprehension posttest had an acceptable level of internal consistency, as determined by a Cronbach's alpha of .805. Participants worked at their own pace and had time as needed to complete the whole task, but on average they spent about 40 min to complete the reading and visual mapping tasks.



Fig. 1 Example of a student visual map

Open-ended map scoring

This present investigation used an open-ended visual information mapping approach to represent participants' text comprehension (i.e., to represent their KS of the text). The rationale is that open-ended mapping will likely obtain richer knowledge elicitation, especially of unanticipated but perhaps important concepts (e.g., most salient to the participants). However, compared to close-ended mapping, the representation and analysis phases for open-ended mapping are more likely to be brittle and thus demand extensive consideration, especially regarding whether important terms are included or not [i.e., a "latent variable" problem, Wilks et al. (2005)].

Using a node degree vector analysis described by Clariana et al. (2015), this current investigation analyzed the open-ended maps based on the 26 terms from the two text conditions including 4 headings phrases, the 7 NIS terms, the 7 SIS terms, and the most frequently used 8 unsigaled terms. To establish a benchmark referent map, three subject domain experts (in L1) negotiated together to create a single referent maps using the selected 26 terms. This referent map was used for comparison to participants' maps.

Results

The data for analysis consists of the comprehension posttest scores and the eighty-eight open-ended individual maps. First, the comprehension multiple-choice posttest data are presented, then following the analyses approach used by Clariana et al. (2015), the map data are described and then compared in four ways including the *term-related measures*: (1) term occurrence and (2) node degree, and the *form-related measures*: (3) graph centrality and (4) pattern-matching measures as correlations with the expert map.

Comprehension posttest

The descriptive statistics of the participants’ multiple-choice comprehension posttest performance are presented in Table 2. A two-way ANOVA was conducted to examine the effects of two factors, text signal (NIS or SIS) and proficiency level (Low or High), on the comprehension posttest. Residual analysis was performed to test for the assumptions of the two-way ANOVA. Outliers were assessed by inspection of a boxplot, normality was assessed using Skewness and Kurtosis normality test for each cell of the design, and homogeneity of variances was assessed by Levene’s test. There were no outliers, residuals were normally distributed (−0.173 for Skewness, 0.512 for Kurtosis) and there was homogeneity of variances ($p = .071$).

The two-way ANOVA showed significant effect for the text signal, $F(1, 52) = 189.41$, $p = .00$, partial $\eta^2 = .88$; for proficiency levels, $F(1, 52) = 16.06$, $p = .01$, partial $\eta^2 = .57$; for the interaction of text signal and proficiency levels, $F(2, 52) = 7.32$, $p = .03$, partial $\eta^2 = .32$. The significant interaction is shown in Fig. 2. Cohen effect sizes are, $d = \text{Low-NIS} (.16) < \text{High-NIS} (.56) < \text{Low-SIS} (.74) < \text{High-SIS} (1.37)$. Note that the low proficient SIS participants (Low-SIS) outperformed the high proficient NIS participants (High-NIS).

Descriptive analysis

Using the *Map-Reader* software tool (contact the authors for access to this tool), we automatically analyzed the “size” of the individual maps as the average total number of conceptual terms (node), including all terms in all maps, not just the selected signal terms. The results are summarized in Table 3. What is surprising is that the Low-SIS group maps were approximately equivalent with the High-NIS maps in terms of quantity (17.8 vs. 18.1) but not in terms of quality, the Low-SIS map term agreement with expert was 47 % compared to 39 % for the High-NIS maps. This unexpected relative advantage for the low-English proficiency participants in the SIS treatment will be further considered in the analyses below.

The degree of a node is the number of links connected to the node. The average node degree is calculated as the average number of links divided by the number of nodes in a network (Clariana et al. 2013); values from 0 to 0.7 indicate an unconnected network while values above 2.0 indicate a highly relational and complex-connected network. Again, the average node degree values for both the Low-SIS and High-NIS groups were approximately equivalent in terms of quantity (1.4 vs. 1.23) but not in terms of quality, the Low-SIS maps had more qualified node degree values compared to the High-NIS maps (41 % vs. 31 % agreement with the expert). That is, the Low-SIS maps had a more well-connected relational and complex map structure dominated by important relevant nodes relative to the High-NIS maps.

Table 2 Mean and SD for comprehension multiple-choice posttest with Cohen’s effect size d (using pooled standard deviation)

Proficiency	Text signals		Mean	d
	NIS	SIS		
Low	2.5 (1.08)	4.9 (1.69)	3.7 (1.19)	1.01
High	4.1 (1.52)	6.7 (2.01)	5.4 (2.04)	
Mean	3.3 (1.77)	5.8 (2.00)		
Effect size, d	1.3			

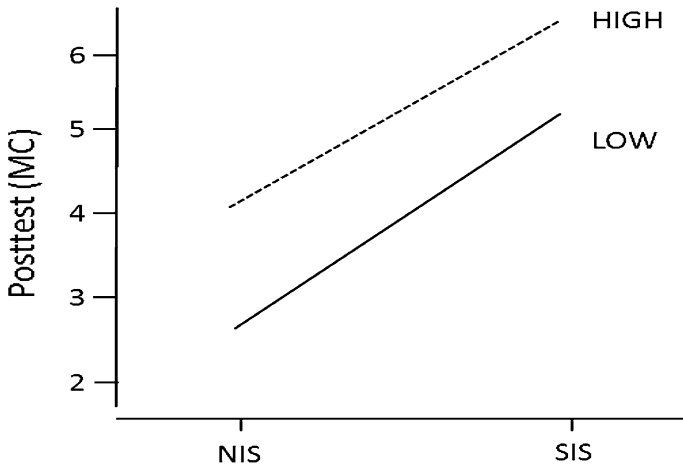


Fig. 2 The significant interaction of text signal and proficiency levels

Table 3 The average number of terms (occurrence) and average node degree data for each condition by proficiency level

	Term occurrence		Mean	Node degree		Mean
	Low	High		Low	High	
NIS	11.3 (Exp. 23 %)	18.1 (Exp. 39 %)	14.7 (Exp. 31 %)	1.35 (Exp. 19 %)	1.23 (Exp. 31 %)	1.29 (Exp. 25 %)
SIS	17.8 (Exp. 47 %)	22.3 (Exp. 64 %)	20.1 (Exp. 56 %)	1.4 (Exp. 41 %)	1.5 (Exp. 61 %)	1.45 (Exp. 51 %)
Mean	14.6 (Exp. 35 %)	20.2 (Exp. 52 %)		1.37 (Exp. 30 %)	1.36 (Exp. 46 %)	

Agreement with the expert is shown in parenthesis (as *Exp. %*)

Prior to the MANOVA presented below, Box's *M* test of equality of covariance matrices was used to check for homogeneity assumption for occurrence data and for node degree data. Box's *M* tests were not significant, $p = .319$ for occurrence data and $p = .248$ for node degree data, suggesting no difference between these variances. Levene's *F* tests of equality of variance matrices were also all not significant for occurrence data (p ranged from .159 to .854), and for node degree data (p ranged from .248 to .942) suggesting the assumption of equal variances was met. The next section reports four separate analyses of the map measures of KS including term-related data (term occurrence and node degree data) and form-related data (graph centrality and pattern matching measures of convergence).

Analysis of map term occurrence data

All maps were converted to 26-element *term occurrence* vectors with a "1" when a term is present in a map and a "0" when absent. Analysis consists of a 1-between, 1-within mixed

MANOVA with the between subjects factor text condition (NIS or SIS) and the within subjects factor term occurrence (as headers, non-important subtopics, important subtopics, and unsigned high frequency terms). The between subjects factor was significant, $F(1, 15) = 23.92, p = .001, \text{partial } \eta^2 = .425$, for the low proficiency condition, and also significant, $F(1, 22) = 42.55, p = .000, \text{partial } \eta^2 = .598$, for the high proficiency condition.

Follow-up ANOVA were conducted for each of these four kinds of terms, for the low proficiency condition, two of these terms were significant ($\alpha = .05$; *Non-important subtopic*, $F(1,18) = 19.00, p = .001, \text{partial } \eta^2 = .724$, and *Important subtopic*, $F(1, 18) = 5.586, p = .034, \text{partial } \eta^2 = .441$). The non-important subtopic terms occurred more in the Low-NIS group maps (63 vs. 34 % of the Low-SIS) and the important subtopic terms occurred more in the Low-SIS group maps (58 vs. 27 % of the Low-NIS). For the high proficiency condition, three of these four terms were significant ($\alpha = .05$; *Header*, $F(1, 25) = 11.04, p = .000, \text{partial } \eta^2 = .644$, and *Non-important subtopic*, $F(1, 25) = 15.28, p = .039, \text{partial } \eta^2 = .570$, and *Important subtopic*, $F(1, 25) = 10.79, p = .017, \text{partial } \eta^2 = .524$). The heading terms occurred more in the High-SIS group maps (76 vs. 60 % of the High-NIS), the important subtopic terms occurred more in the High-SIS maps (81 vs. 52 % of the High-NIS), and the non-important subtopic terms occurred more in the High-NIS group maps (80 vs. 43 % of the High-SIS; see Table 4; Fig. 3).

Analysis of map node degree data

All maps were converted to 26-element *node degree* vectors with the number of links to each node. Following Clariana et al. (2015), unlinked nodes on a map were given a node degree value of “1” (self–self link) and any nodes that did not occur on a map were given a node degree of “0”. Using the same approach described for the term occurrence data above, a 1-between, 1-within mixed MANOVA with the between subjects factor text condition (NIS or SIS) and the within subjects factor term node degree was conducted, and the analysis for the low proficiency condition was significant, $F(1, 15) = 15.44, p = .001, \text{partial } \eta^2 = .446$ and was also significant for the high proficiency condition, $F(1, 22) = 32.14, p = .021, \text{partial } \eta^2 = .412$.

Follow up ANOVA were conducted for each of these four kinds of terms, for the low proficiency condition, two of these were significant ($\alpha = .05$; *Non-important subtopic*, $F(1, 18) = 8.78, p = .000, \text{partial } \eta^2 = .957$, and *Important subtopic*, $F(1, 18) = 9.90, p = .000, \text{partial } \eta^2 = .430$). The important subtopic terms predominated in the Low-SIS

Table 4 Average frequency of *occurrence* and standard deviations for the NIS (n = 44; 23 of low proficiency, 21 of high proficiency) and SIS maps (n = 44; 23 of low proficiency, 21 of high proficiency) with Cohen’s effect size *d* (using pooled standard deviation) and significance level (*p*)

Term categories	Low proficiency		<i>d</i>	<i>p</i>	High proficiency		<i>d</i>	<i>p</i>
	NIS	SIS			NIS	SIS		
Header terms	0.60 (.35)	0.64 (.29)	0.50	.45	0.60 (.29)	0.76 (.15)	1.50	.00
Non-important subtopic terms	0.63 (.13)	0.34 (.10)	1.31	.00	0.80 (.23)	0.43 (.19)	0.90	.01
Important subtopic terms	0.27 (.21)	0.58 (.20)	1.00	.03	0.52 (.30)	0.81 (.24)	−1.10	.02
Unsigned terms	0.22 (.38)	0.25 (.31)	1.45	.15	0.45 (.41)	0.48 (.28)	0.35	.15

Values can range from 0 to 1.00; the standard deviations are shown in parenthesis

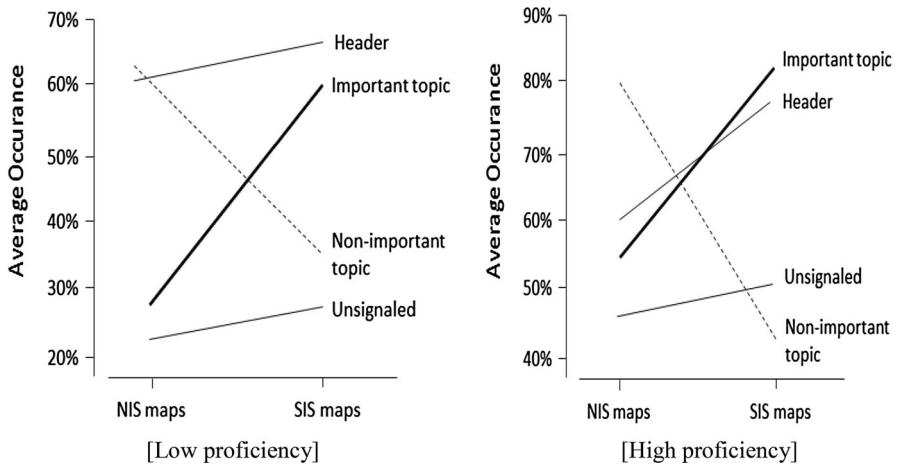


Fig. 3 Map occurrence values for each type of term by proficiency level

representations (1.7 vs. 0.6 of the Low-NIS) while non-important subtopic terms in the Low-NIS (1.9 vs. 0.5 of the Low-SIS). For the high proficiency condition, three of these were significant ($\alpha = .05$; *Header*, $F(1, 25) = 6.51$, $p = .038$, $\text{partial } \eta^2 = .244$, *Non-important subtopic*, $F(1, 25) = 12.53$, $p = .002$, $\text{partial } \eta^2 = .555$, and *Important subtopic*, $F(1, 25) = 3.50$, $p = .015$, $\text{partial } \eta^2 = .324$). Both the headings (2.3 vs. 1.0) and important subtopic terms (2.0 vs. 1.2) predominated in the High-SIS representations while the non-important subtopic terms (1.7 vs. 0.8) in the High-NIS (see Table 5; Fig. 4).

Graph centrality data

Using the node degree vectors above for the 26 terms, graph centrality was calculated as a holistic measure of network form, or structure, that ranges from 0 (linear form) to 1 (star form), with mid-range values (0.4–0.6) indicating optimally relational and complex network form (Clariana et al. 2011; Kim and Clariana 2015). The graph centrality data were analyzed by ANOVA, with the factor text condition (NIS vs. SIS) by proficiency level. The results are summarized in Table 6.

Table 5 Average *node degree* and standard deviation for the NIS ($n = 46$) and SIS maps ($n = 42$) with cohen's effect size d (using pooled standard deviation) and significance (p)

Term categories	Low proficiency		d	p	High proficiency		d	p
	NIS	SIS			NIS	SIS		
Header terms	2.0 (.41)	2.2 (.25)	1.42	.35	1.0 (.29)	2.3 (.15)	1.40	.00
Non-important subtopic terms	1.9 (.19)	0.5 (.11)	1.60	.00	1.7 (.23)	0.8 (.19)	1.00	.01
Important subtopic terms	0.6 (.23)	1.7 (.20)	1.20	.01	1.2 (.30)	2.0 (.24)	-0.80	.02
Unsignaled terms	0.9 (.21)	1.0 (.49)	0.40	.16	0.8 (.41)	1.3 (.28)	1.35	.15

Values can range from 0 to 5.00; the standard deviations are shown in parenthesis

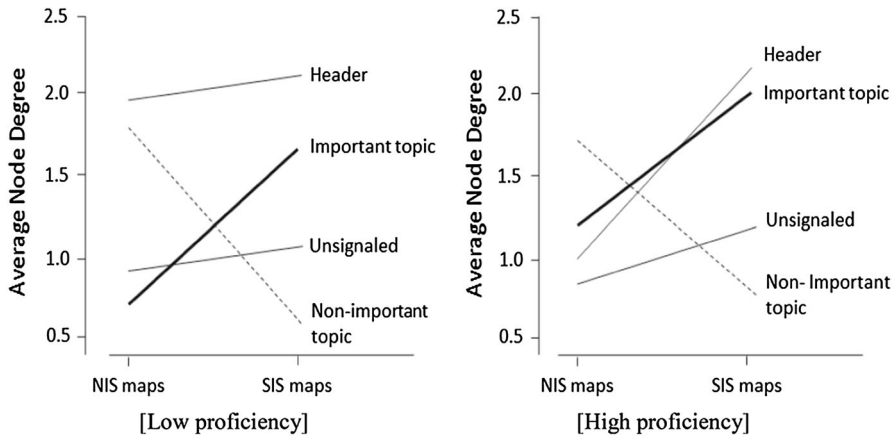


Fig. 4 Map node degree values for each type of term by proficiency level

Table 6 Average *graph centrality* for the NIS (n = 46) and SIS maps (n = 42) with cohen’s effect size *d* (using pooled standard deviation) and significance (*p*)

	Graph centrality		<i>d</i>	<i>p</i>
	NIS	SIS		
Low proficiency	.21 <i>linear</i> (.18)	.43 <i>relational</i> (.29)	1.5	.01
High proficiency	.34 <i>linear</i> (.25)	.49 <i>relational</i> (.11)	1.8	.01

Expert’s $C_{graph} = 0.47$ (*relational*)

The most interesting finding was that the Low-SIS maps were a network-like *relational* structure relative to the High-NIS maps that had a more *linear* structure ($C_{graph} = .43$ vs. $.34$). This finding could be explained in part by those of term-related measures above, reporting that the Low-SIS maps had more terms and more of the important terms (term occurrence, see Table 4) and also the Low-SIS maps were more dominated by important subtopic terms (node degree, see Table 5) relative to the High-NIS maps.

Our working assumption is that a proper relational KS would be better able to support inferences (as gist knowledge) from the text compared to a linear KS that would be better for verbatim knowledge tasks. To consider this assumption, all individual map graph centrality vectors were compared to their *global inferences* comprehension posttest scores, that are part of the *The Cave of Lascaux* TOEFL reading passage. Note that global inferences would require an appropriately structured situation model (i.e., a relational-hierarchical representation) while local inferences can be usually answered employing just the text surface structure representation (i.e., a sequential-linear representation).

The High-SIS, Low-SIS, and High-NIS maps shown in Fig. 5 all show a significant curvilinear relationship between graph centrality and posttest comprehension performance ($r^2 = .39, .28, .31$ respectively). Interestingly, a non-linear “inverted U” relationship was observed in all SIS groups (High-SIS & Low-SIS) suggesting an optimal KS structure. In other words, perhaps too little structure (i.e., a deficient map) or too much structure (i.e., non-appropriated structure dominated by irrelevant terms) both negatively affect

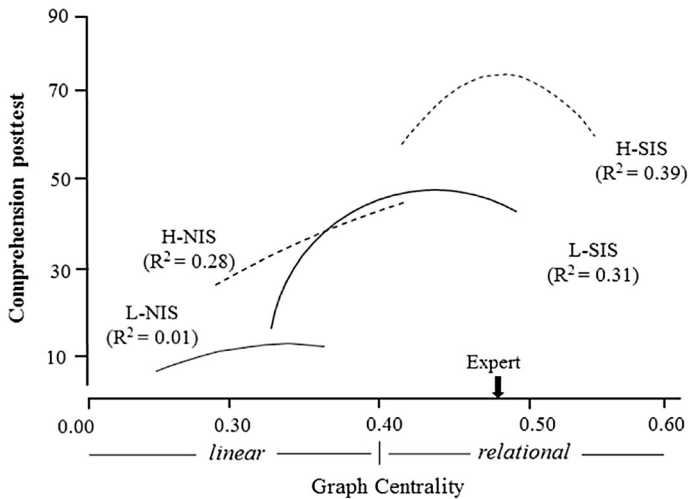


Fig. 5 The relationship between graph centrality and performance on the comprehension posttest. *Dashed line* high proficiency (H), *Solid line* low proficiency (L)

performance on the text comprehension posttest, but the optimal structure (Range of $C_{graph} = .40-.50$; expert's $C_{graph} = .47$) relates to the highest posttest scores. The implication is that specific content can take one or a few patterns within the bounded framework or ontological conceptual space; heuristically stated, “form relates to function”.

Pattern matching measures of convergence

Using node degree vectors used above to calculate graph centrality, each individual's map vector was compared to the expert's vector and to other participants' vectors using Pearson correlation (i.e., knowledge convergence; see Table 7). Note that since these correlation values (r) are not additive (because Pearson r is not interval-level data), then all correlation values were converted into *Fisher z* values (z) using the *MS excel Fisher z* function prior to averaging and statistical comparison of the groups. But since *Fisher z* values are not commonly reported, these values were reported along with the map vector percent overlap—

Table 7 Average correlation as *Fisher z* and average percent overlap of map node degree vectors to the expert and to other maps for the NIS and SIS maps with cohen's effect size d (using pooled standard deviation) and significance (p)

	Low proficiency				d	p	High proficiency				d	p
	NIS		SIS				NIS		SIS			
Expert to	.30 (.11)	8 %	.63 (.19)	31 %	0.98	.02	.60 (.19)	29 %	.93 (.18)	53 %	1.28	.00
NIS maps to	.31 (.13)	9 %	.32 (.28)	10 %	0.01	.87	.65 (.23)	32 %	.40 (.29)	14 %	0.80	.04
SIS maps to	.32 (.20)	10 %	.85 (.24)	48 %	1.00	.03	.40 (.31)	14 %	1.13 (.34)	66 %	1.00	.01

Standard deviations shown in parenthesis

similarity *estimate* that is easier to understand. To represent the percent overlap, the *Fisher z* values were transformed to an *r* correlation value using the *Fisher z* inverse function in *MS Excel* and then this *r* value was squared into coefficient of determination (r^2).

Regarding *map vector similarity to the expert*, for the low proficiency participants, ANOVA of their *Fisher z* values was significant, $F(1, 18) = 42.232, p = .021$, the Low-SIS maps converged more with the expert (31 % overlap, *Fisher z* = .63) compared to the Low-NIS maps with the expert (8 % overlap, *Fisher z* = .30). For the high proficiency, ANOVA of their *Fisher z* values was also significant, $F(1, 140) = .39, p = .001$, the High-SIS maps converged more with the expert (53 % overlap, *Fisher z* = .93) relative to the High-NIS maps with the expert (29 % overlap, *Fisher z* = .60).

Regarding *map to map similarity*, for the low proficient, ANOVA of their *Fisher z* values was significant, $F(1, 25) = 12.563, p = .032$, the Low-SIS maps converged with each other (48 % overlap, *Fisher z* = .85) compared to the Low-NIS maps with each other (9 % overlap, *Fisher z* = .31). For the high proficiency, ANOVA of their *Fisher z* values was also significant, $F(1, 20) = 9.78, p = .013$, the High-SIS maps converged more with each other (66 % overlap, *Fisher z* = 1.13) relative to the High-NIS maps with each other (32 % overlap, *Fisher z* = .65).

In summary, the SIS maps in both proficiency conditions were more like the expert map and more like each other (i.e., had a homogenous KS) perhaps due to the coherent underlined important subtopic structure across the entire text, while the NIS maps were a little more idiosyncratic perhaps due to the incoherent underlined non-important subtopic structure across the entire text although the High-NIS maps were somewhat alike (32 % overlap). But, this unexpected convergence of the High-NIS map structure was consistent with that of Clariana et al. (2015) investigation, who interpreted the text structure similarity of the underlined non-important subtopic version (in their case incidental hyperlinked version) as indicating that readers tend to rely on a “linear/list reading strategy” in which the text is viewed as a list/collection of loosely linked concepts or terms in the text (Meyer et al. 2012) when the text is not well comprehended by the readers.

Replicating Clariana et al. (2015) analysis approach to consider this linear/list strategy conjecture, we numbered the signaled 18 terms (except the 8 unsignaled terms) based on their serial order of first occurrence in the text, and then conducted Pearson correlation with the High-NIS and High-SIS average term occurrence vectors (also except the 8 unsignaled terms); the rank order of first occurrence of signaled terms in the text is related to the term occurrence vectors in both the High-NIS ($r = .48$) and High-SIS ($r = 0.21$) maps; the initially mentioned signal terms in the text had the greater average occurrence (a primary serial position effect for these bilingual-created artifacts), and the influence of this linear sequence was much more obvious in the High-NIS group [$r = .48$ (23 %) vs. $r = .21$ (4 %)]. Future research should consider the possibility that bilingual readers’ KS converges on the sequentially signaled terms, when the L2 text is too difficult to be well apprehended by bilingual readers.

Conclusions

This investigation explored the influence of text signals in a print-based text passage on L2 science expository text comprehension by proficiency level. The results completely replicated the patterns observed by Clariana et al. (2015) study in L1 reading, adding substantial generalizability for the influence of text signals on KS.

For the low proficiency readers

Their term-related data show that the Low-NIS maps were dominated by the headings and non-important subtopic terms while the Low-SIS maps were dominated by the headings and important subtopic terms they read (see Figs. 3, 4). This finding indicates that the low proficient readers' maps were strongly influenced by the text signals they read, which means that they strongly depend on the text signals for their L2 expository text comprehension. Further, the Low-SIS maps relative to the Low-NIS maps on average were more *relational* in form ($C_{graph} = 0.43$ vs. 0.21) and their relational form became more alike (48 vs. 9 %) and more like the expert (31 vs. 8 %), which relates to higher comprehension posttest performance (4.9 vs. 2.5).

For the high proficiency readers

The term-related data show that the High-SIS maps were strongly dominated by the headings and important subtopic terms while the High-NIS maps were dominated by only non-important subtopic terms they read, not by headings (see Figs. 3, 4). The term data indicate that the signaled important subtopic terms in the SIS text led the high proficient readers to integrate the important subtopics and the four headings relatively more frequently (the greater term occurrence, Table 4) and more centrally into their maps (the greater node degree, Table 5), while the non-important subtopics in the NIS text did not. Further, the High-SIS maps relative to the High-NIS maps on average were *relational* in form ($C_{graph} = 0.49$ vs. 0.34), and their relational maps converged substantially more with each other (66 vs. 32 %) and with the expert (53 vs. 29 %), which relates to higher comprehension posttest performance (6.7 vs. 4.1).

Discussion

These findings from the term- and form-related data in both proficiency levels indicate that the NIS and SIS maps (as KS) were fundamentally different, suggesting that attending to specific text terms while reading strongly influenced the L2 readers' term use (measured as term occurrence) and importance of the terms (measured as node degree) and also the organization of the maps, KS (measured as graph centrality). These results line up with those of previous studies. Previous studies (in monolingual, native language) have clearly demonstrated that even though the signaled text *misrepresented* the topic structure of the text, the readers were very heavily influenced by the signals, arguing that signals, either text topic or non-text topic, affect how readers represent a text (Meyer and Rice 1983; Lorch and Lorch 1995). Subsequent investigations have shown that text signals lead to different processing strategies ['signaled-guided processing', Lorch et al. (2001)] which can, in turn, result in readers' different representations of the text's topic structure and thus their comprehension of the text (Meyer et al. 2012; Ritchey et al. 2008).

This strong 'signaled-guided processing' might account for why the Low-SIS text readers perform better than the High-NIS text readers. Given that the SIS text has the coherent/logical connection between the headings and underlined important subtopic signaled terms, the SIS text might help the low proficient L2 readers in establishing a structured/relational representation of the text's topic structure by directing the readers to the coherent text signals and their organization [i.e., "structure reading strategy" using top-down processing, see Meyer et al. (2012)]; whereas, the NIS text has the less coherent/

Table 8 Possible reading strategy, term data, map (KS) form data, and posttest performance by proficiency levels

	Text read	Reading strategy	Terms dominated	Map(KS) form	Posttest
Low	NIS text	List strategy (bottom-up processing)	Headings	NIS Linear	Low
	SIS text	Structure strategy (top-down processing)	Headings	SIS Relational	Intermediate
High	NIS text	List strategy (bottom-up processing)		NIS Linear	Intermediate
	SIS text	Structure strategy (top-down processing)	Headings	SIS Highly relational	High

illogical connection between the headings and underlined non-important subtopic signaled terms, which might interfere the high proficient L2 readers in building the relational/hierarchical organization, but instead, lead them to arrive at only a collection of loosely related concepts or terms in the text [i.e., “linear/list reading strategy” using bottom-up processing, see Meyer et al. (2012)] or to somewhat disregard the headings (see Figs. 3, 4); Clariana et al. (2015) study in L1 expository text reading observed a similar pattern where headings were disregarded to some extent when headings and signaled terms (incidental hyperlinks) are not coherently related in the text.

Thus, this investigation suggests that reading texts with coherent text signals (SIS condition) substantially improves the qualities of the underlying mental structure related to L2 text comprehension as measured through the map artifacts, even of the low proficient L2 readers, and whereas, reading texts with less coherent text signals (the NIS condition) debases the qualities of the mental structure and L2 artifact of bilingual readers, even of high proficient L2 readers, and this suboptimal KS is relatively less related to L2 text comprehension. Also surprisingly the high proficient participants were less likely to integrate the headings into their maps (i.e., KS), a finding not previously reported in L2 reading but that support the findings of Clariana et al. (2015) in L1 reading (see Table 8). Future research should consider the relationship between headings and underlined terms in L2 reading and how the relationship influences L2 reading comprehension.

Implications and limitations

Although it is not surprising that the underlined non-important subtopic terms or important subtopic terms were included far more often in the maps of those who received those signals, the findings have an important practical application. The coherent text signals used here (i.e., headings and underlined important subtopic terms) appear to establish appropriate frameworks or ontological conceptual spaces for the L2 expository text content, even for the low proficient L2 readers, by allowing them to more readily identify the text’s overall topic structure, or the relationship between the main ideas in the text. As ‘better’ association networks are established in memory, then top-down processing of this content (i.e., structural reading strategy) will be facilitated, and the bilinguals will be more able to read unsigned domain-related texts. Thus practically speaking, instructional designers must place a higher priority on text signals during the analysis, design, and development of learning materials; for example, by explicitly establishing and then verifying that signals in the materials align with the actual domain-normative knowledge structure. Further, instructors should pay attention to the text signals in assigned lesson materials to confirm

that these support the desired learning outcomes. Such a focus on text signals requires a relatively small time commitment by designers and teachers but obtains a fairly substantive improvement in learning.

From a broader theoretical perspective, the structure inherent in lesson materials and artifacts, both the explicit content as well as the format of that content, influences or is even 'imprinted' onto the learner's knowledge structure at least in the short term and this influences cognition. In this view, an author's knowledge structure is imbued to the reader through the text artifacts, structure is what is passed from person to person, and experts' knowledge structure converges as domain-normative knowledge. Thus accounting for knowledge structure in people and in artifacts has theoretical and practical implications.

This investigation could be criticized that it would be unusual or even counterproductive to create a text like the NIS condition that signals subtopics that although may be interesting, are not actually central to the text topic structure. In the previous investigation by Clariana et al. (2015), hyperlinks in a Wikipedia document were shown to be processed by the readers as though these were intentional text signals and not as incidental links to other Wikipedia articles. Many style sheets have since removed underlining as a hyperlink signal and simply use light blue (this style change occurred in Wikipedia about 2004), although Microsoft Word © still utilizes this convention. More to the point, modern textbooks use text signals profusely to indicate both text topic structure AND just interesting content (e.g., a text signal and sidebar on General Custer's horse Comanche in a high school history textbook). Further, social annotation is a current area of research that allows multiple readers of an online document to highlight and annotate the document that is then seen by subsequent readers [e.g., see Li et al. (2015)]. The findings of this current investigation would suggest that during such social annotation, the text signals left by the first readers would profoundly influence the later readers [privilege of position, Gernsbacher's (1991), Structure Building Framework]. So analogs of the NIS condition used here may regularly occur *in vivo*, and so these findings reported here bear on these kinds of texts.

Another limitation is the lack of random assignment. The TOEFL data and the nature of the two sections support pre-intervention equivalence and so these findings are likely due to the intervention. Thus, these findings should not be over-generalized. Further, the mapping task completed by all participants preceded the comprehension posttest and so mapping may have influenced the test performance. Possibly, the findings reported here for NIS an SIS at different proficiency levels may or may not be observed if the mapping task is not used.

But what are the implications for lesson texts where no terms are underlined? We can't be sure, but we propose that eye-tracking saccades of un signaled text would follow quite person-specific idiosyncratic patterns, imagine it as self-directed implicit underling of terms. The high proficient bilinguals are more likely to attend to more of the SIS terms through top-down processing of the text, while low proficient bilinguals will attend to terms far more randomly due to idiosyncratic familiarity with the terms. The quality of the association networks established in memory would reflect the attentional sequence. Future research should compare a SIS signaled to a non-signaled text condition to consider this hypothesis that low proficient bilinguals will establish nearly random KS in the non-signaled condition while the high proficient bilinguals in the nonsignaled condition will establish a coherent KS more related to the text's topic structure; and that both low and high proficient would benefit from proper signals.

In addition, these findings strongly support those of previous investigations of the validity and reliability of this mapping and analysis approach as a measure of KS [e.g.,

Kim (2012), Pirnay-Dummer and Ifenthaler (2010), Villalon and Calvo (2011)] that can be applied in second language settings (Kim and Clariana 2015). The term occurrence and node degree data almost exactly mirror the findings of Clariana et al. (2015) even though the participants and the lesson materials were radically different. The data clearly show that KS mediated comprehension posttest performance. In conclusion, KS measures are useful in both monolingual and also multi-language settings.

Acknowledgments I thank my advisor and co-author of this paper, Dr. Roy B. Clariana, for his assistance and comments that greatly improved the manuscript (You are the best ever advisor). I also thank “anonymous” reviewers for their so-called insights.

Compliance with ethical standards

Conflict of Interest The authors declare that they have no conflict of interest.

Appendix

The Cave of Lascaux text passage with the NIS terms shown within brackets and SIS terms shown with underlines>.

Lascaux Cave Paintings

In Southwest France in the 1940s, playing children discovered Lascaux Grotto, a series of [narrow cave chambers] that contain *huge prehistoric paintings of animals*. Many of these beasts are as large as 16 feet (almost 5 meters). Some follow each other in solemn parades, but others swirl about, sideways and upside down. The animals are bulls, wild horses, reindeer, bison, and mammoths outlined with charcoal and painted mostly in reds, yellow, and browns. Scientific analysis reveals that the colors were derived from ocher and other iron oxides ground into a fine powder. Methods of applying color varied: some colors were brushed or smeared on rock surfaces and others were blown or sprayed. It is possible that [tubes made from animal bones] were used for spraying because hollow bones, some stained with pigment, have been found nearby.

One of the most *puzzling aspects of the paintings is their location*. Other rock paintings—for example, those of Bushmen in South Africa—are either located [near cave entrances or completely in the open]. Cave paintings in France and Spain, however, are in recesses and caverns far removed from original cave entrances. This means that artists were forced to work in cramped spaces and without sources of natural light. It also implies that whoever made them did not want them to be easily found. Since [cave dwellers normally lived close to entrances], there must have been some reason why so many generations of Lascaux cave dwellers hid their art.

Migration Opinion

Scholars offer three related but different opinions about the mysterious origin and significance of these paintings. One opinion is *that the paintings were a record of seasonal migrations* made by herds. Because some paintings were made directly over others *overpainting*, obliterating them, it is probable that a painting’s value ended with the migration it pictured. Unfortunately, this explanation fails to explain the hidden locations, unless the migrations were celebrated with [secret ceremonies].

Hunting Opinion

Another opinion is that the paintings were directly related to hunting and were an essential part of a special preparation ceremony. This opinion holds that the pictures and whatever ceremony they accompanied were *an ancient method of psychologically motivating hunters*. It is conceivable that before going hunting the hunters would draw or study pictures of animals and imagine a successful hunt. Considerable support exists for this opinion because several animals in the pictures are wounded by arrows and spears. This opinion also *attempts to solve the overpainting* by explaining that an animal's picture had no further use after the hunt.

Ceremonial Opinion

A third opinion takes psychological motivation much further into the realm of tribal ceremonies and mystery: the belief *that certain animals assumed mythical significance* as ancient ancestors or protectors of a given tribe or clan. Two types of images substantiate this theory: the strange, indecipherable geometric shapes that appear near some animals, and the few drawings of men. Wherever men appear they are crudely drawn and their bodies are elongated and rigid. Some men are in a prone position and some have bird or animal heads. Advocates for this opinion point to reports from people who have experienced a trance state, a highly suggestive state of low consciousness between waking and sleeping. Uniformly, these people experienced weightlessness and the sensation that their bodies were being stretched lengthwise. Advocates also point to people who believe that the forces of nature are inhabited by spirits, particularly *shamans* who believe that an animal's spirit and energy is transferred to them while in a trance. One Lascaux narrative picture, which shows [a man with a birdlike head] and a wounded animal, would seem to lend credence to this third opinion, but there is still much that remains unexplained. For example, where is the proof that the man in the picture is a shaman? He could as easily just be a hunter wearing a headmask. Many tribal hunters, including some (Native Americans), camouflaged themselves by wearing animal heads and hides.

Perhaps so much time has passed that there will never be satisfactory answers to the cave images, but their mystique only adds to their importance. Certainly a great art exists, and by its existence reveals that ancient human beings were not without intelligence, skill, and sensitivity.

The four headings

Lascaux cave paintings, Migration opinion, Hunting opinion, Ceremonial opinion.

The seven non-important subtopic terms (NIS)

Narrow cave, tubes, near entrances, cave dwellers, secret ceremonies, birdlike head, and Native Americans.

The seven important subtopic terms (SIS)

Paintings of animals, puzzling, location, overpainting, motivation, mystical significance, and shamans.

The eight unsigaled terms

Intelligent, ancient human, arrow and spear, trance state, France and Spain, children, and hidden (in order from high to low frequency).

References

- Barry, S., & Lazarte, A. (2000). A comparison of first and second language readers' situation models as evidenced in a recall task. (ERIC Document Reproduction Service No. ED 441 219) <http://files.eric.ed.gov/fulltext/ED441219.pdf>.
- Cashen, V. M., & Leicht, K. L. (1970). Role of the isolation effect in a formal educational setting. *Journal of Educational Psychology*, *61*(6p1), 484.
- Clariana, R. B. (2010). Deriving group knowledge structure from semantic maps and from essays. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 117–130). New York: Springer.
- Clariana, R. B., Draper, D., & Land, S. M. (2011). An automated measure of group knowledge structure convergence. In *annual meeting of the AECT International Convention*. Retrieved September (Vol. 22, p. 2014).
- Clariana, R. B., Engelmann, T., & Yu, W. (2013). Using centrality of concept maps as a measure of problem space states in computer-supported collaborative problem solving. *Educational Technology Research and Development*, *61*(3), 423–442.
- Clariana, R. B., Rysavy, M. D., & Taricani, E. (2015). Text signals influence team artifacts. *Educational Technology Research and Development*, *63*(1), 35–52.
- Clariana, R. B., Wolfe, M. B., & Kim, K. (2014). The influence of narrative and expository lesson text structures on knowledge structures: Alternate measures of knowledge structure. *Educational Technology Research and Development*, *62*(5), 601–616.
- Crouse, J. H., & Idstein, P. (1972). Effects of encoding cues on prose learning. *Journal of Educational Psychology*, *63*(4), 309.
- Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing*, *15*(7–8), 739–757.
- Fecteau, M. L. (1999). First-and second-language reading comprehension of literary texts. *The Modern Language Journal*, *83*(4), 475–493.
- Fesel, S. S., Segers, E., Clariana, R. B., & Verhoeven, L. (2015). Quality of children's knowledge representations in digital text comprehension: Evidence from pathfinder networks. *Computers in Human Behavior*, *48*, 135–146.
- Fowler, R. L., & Barker, A. S. (1974). Effectiveness of highlighting for retention of text material. *Journal of Applied Psychology*, *59*(3), 358.
- Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 27, pp. 217–263). San Diego: Academic Press.
- Hartley, J., Bartlett, S., & Branthwaite, A. (1980). Underlining can make a difference—sometimes. *The Journal of Educational Research*, *73*(4), 218–224.
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in second language acquisition*, *24*(04), 541–577.
- Jonassen, D. H. (1985). *The technology of text: Principles for structuring, designing, and displaying text* (Vol. 2). Englewood Cliffs: Educational Technology.
- Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kim, M. (2012). Cross-validation study on methods and technologies to assess mental models in a complex problem solving situation. *Computers in Human Behavior*, *28*(2), 703–717.
- Kim, K., & Clariana, R. B. (2015). Knowledge structure measures of reader's situation models across languages: Translation engenders richer structure. *Technology, Knowledge and Learning*, *20*, 249–268.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction- integration model. *Psychological Review*, *95*(2), 163–182.
- Laborda, J. G. (2009). Building a validity argument for the Test of English as a Foreign Language™. *ELT Journal*, *63*(3), 291–294.

- LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research*, 20, 75–91.
- Lambiotte, J. G., Dansereau, D. F., Cross, D. R., & Reynolds, S. B. (1989). Multi-relational semantic maps. *Educational Psychology Review*, 1(4), 331–367.
- Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, 57(1), 87–118.
- Lee, J. W., & Schallert, D. L. (1997). The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly*, 31(4), 713–739.
- Lemarié, J., Lorch, J. R. F., & Péry-Woodley, M. P. (2012). Understanding how headings influence text processing. *Discours. Revue de linguistique, psycholinguistique et informatique*, (10).
- Leow, R. P. (2001). Do learners notice enhanced forms while interacting with the L2?: An online and offline study of the role of written input enhancement in L2 reading. *Hispania*, 84, 496–509.
- Li, S. C., Pow, J. W. C., & Cheung, W. C. (2015). A delineation of the cognitive processes manifested in a social annotation environment. *Journal of Computer Assisted Learning*, 31, 1–13.
- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(03), 673–680.
- Lorch, R., Lemarié, J., & Grant, R. (2011). Signaling hierarchical and sequential organization in expository text. *Scientific Studies of Reading*, 15(3), 267–284.
- Lorch, R. F., Jr., & Lorch, E. P. (1995). Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology*, 87(4), 537.
- Lorch, R. F., Lorch, E. P., Ritchey, K., McGovern, L., & Coleman, D. (2001). Effects of headings on text summarization. *Contemporary Educational Psychology*, 26(2), 171–191.
- Meyer, B. J., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, 16, 72–103.
- Meyer, B. J. F., Ray, M. N., & Middlemiss, W. (2012). Children's use of comparative text signals: The relationship between age and comprehension ability. *Discours[Enligne]*. Retrieved from <http://discours.revues.org/8637>.
- Meyer, B. J., & Rice, G. E. (1983). Learning and memory from text across the adult life span. In J. Fine & R. O. Freedle (Eds.), *Developmental studies in discourse* (pp. 291–306). Norwood, NJ: Ablex.
- Nist, S. L., & Hogrebe, M. C. (1987). The role of underlying and annotating in remembering textual information. *Literacy Research and Instruction*, 27(1), 12–25.
- Overstreet, M. H. (2002). The effect of textual enhancement on second language learner reading comprehension and form recognition (Doctoral dissertation, University of Illinois at Urbana-Champaign).
- Perfetti, C. A. (1989). There are generalized abilities and one of them is reading. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 307–336). Hillsdale, NJ: Lawrence Erlbaum.
- Pirnay-Dummer, P., & Ifenthaler, D. (2010). Automated knowledge visualization and assessment. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge*. New York: Springer.
- Ritchey, K., Schuster, J., & Allen, J. (2008). How the relationship between text and headings influences readers' memory. *Contemporary Educational Psychology*, 33(4), 859–874.
- Ryan, C. (2013). American Community Survey Reports: Language use in the United States: 2011. Retrieved from <https://www.census.gov/prod/2013pubs/acs-22.pdf>.
- Shook, D. J. (1994). FL/L2 Reading, grammatical information, and the input-to-intake phenomenon. *Applied Language Learning*, 5(2), 57–93.
- Spector, J. M., Johnson, T. E., & Young, P. A. (2015). An editorial on replication studies and scaling up efforts. *Education Technology Research and Development*, 63, 1–4.
- Spyridakis, J. H., & Standal, T. C. (1987). Signals in expository prose: Effects on reading comprehension. *Reading Research Quarterly*, 22(3), 285–298.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension*. New York: Academic Press.
- van Hell, J. G., & Kroll, J. F. (2013). Using electrophysiological measures to track the mapping of words to concepts in the bilingual brain: A focus on translation. In J. Altarriba & L. Isurin (Eds.), *Memory, language, and bilingualism: Theoretical and applied approaches*. New York: Cambridge University Press.
- Villalon, J., & Calvo, R. A. (2011). Concept maps as cognitive visualizations of writing assignments. *Educational Technology & Society*, 14(3), 16–27.
- Wilks, C., Meara, P., & Wolter, B. (2005). A further note on simulating word association behaviour in a second language. *Second Language Research*, 21(4), 359–372.

- Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13*, 109–132.
- Zareva, A. (2007). Structure of the second language mental lexicon: how does it compare to native speakers' lexical organization? *Second Language Research, 23*(2), 123–153.
- Zareva, A., & Wolter, B. (2012). The 'promise' of three methods of word association analysis to L2 lexical research. *Second Language Research, 28*(1), 41–67.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162.

Kyung Kim is a Doctoral student in Learning, Design, and Technology program at the Pennsylvania State University. He has worked with his advisor, Dr. Roy Clariana, on the measures and applications of knowledge structure in various learning environments.

Roy B. Clariana is a Professor of Learning, Design, and Technology program and Department Head of Learning and Performance Systems at the Pennsylvania State University. His current research focus is on knowledge structure measures.