Exploring Semantic Discontinuity in Middle School Science Texts: Implications for Science Literacy Development

Understanding the Language of Middle School Science: A comparison of discourse marker usage in science and social studies textbooks

Objectives

In the United States, most teachers rely on science textbooks for their science instruction (Weiss et.al,2003). Therefore, much of the knowledge transfer that occurs in these academic contexts relies on students' ability to use and understand the information presented in those materials. In this context, literacy serves as both a goal in and of itself as well as a tool for the acquisition of further knowledge. The challenge in extracting information from a textbook is rooted not only in the difficulty of confronting and understanding novel concepts, but furthermore in the task of parsing and interpreting the language that is used to convey those concepts.

Our focus in this paper is the use of overt discourse markers in the context of middle school textbooks. Building on the growing interest in developing subject-specific literacy strategies (Jetton&Shanahah,2012), we compare discourse-marker (DM) usage across two subject areas—science and social studies. We ask how often sentences in these two subject areas contain DMs and furthermore what variation exists among the types of discourse relations that are signaled overtly. Our aim is to provide a snapshot of the kind of language that is used in different subject areas.

Theoretical Framework

A large body of literature in linguistics grapples with factors that contribute to the meaning, production, and processing of coherent discourse (Asher & Lascarides,2003; Fraser,2006; Grimes,1975; Halliday&Hasan,1976; Hobbs,1979; Kehler,2002; Mann & Thompson,1988; Nemo,2007; Polanyi,1988). Among those factors, discourse-markers (DMs) have been described as an important element in producing coherent texts (McNamara,2001). According to Fraser (2006), DMs are defined functionally (i.e. rather than syntactically) as lexical elements that signal a semantic relationship that holds between discourse segments.

Fraser (2006) indicates that every discourse marker signals one of four types of relationships, and thus, can be classified as: elaborative, contrastive, inferential, and temporal. These categories are exemplified by the conjunctive elements *and*, *but*, *so*, and *then* respectively (Halliday&Hasan,1976; Fraser,2006). Consider the following excerpt from a science textbook:

1. Asteroids are left over from the formation of the solar system about 4.6 billion years ago. It is thought that they crashed into the inner planets during the early period of our solar system. Asteroids lack enough gravity to have an atmosphere. *Consequently*, their surfaces have many craters from impacts with other objects.
   (*Focus on Physical Science*, Glencoe,2007,p.489)

To assign meaning to (1) as a whole, the reader must establish that the text is coherent—that the sentences do not appear together arbitrarily but, rather, relate to each other in meaningful ways (Halliday&Hasan,1976; Taboada&Mann,2006). To interpret (1) as a coherent passage requires that the last sentence be understood as a description of the state of affairs that arises as a result of asteroids' limited gravity. This relationship is signaled in (1) overtly with the DM *Consequently*, an inferential DM according to Fraser (2006).

The meaning associated with *Consequently* as a DM allows for the intended interpretation of this text: that limited gravity results in a lack of atmosphere which renders asteroids vulnerable to frequent impacts which in turn results in numerous craters on the surface of asteroids. This is evidenced by the alternative interpretation that arises when the marker is changed: “Asteroids lack enough gravity... *because* their surfaces have many craters”, where it is the cratered surface that is identified as the cause of the limited gravity. In fact, a discourse marker not only signals how two clauses are intended to relate, but it can also generate inferences of its own. In the case of (1), the consequence described in the last sentence can only be understood if an additional piece of information is accommodated—that the lack of an atmosphere makes asteroids vulnerable to repeated impact from other objects. This information is not stated directly in the text but must be inferred in order for the passage to make sense. An example like this highlights the pragmatic reasoning that is implicated in the establishment of a coherent discourse, a process in which discourse markers play a key role.

Using Fraser’s DM typology, we aim to better understand the nature of discourse relations in science texts and how this varies with non-science registers.

Data Sources

In order to conduct a large-scale comparison of the use of discourse markers across science and non-science textbooks, we compiled a corpus of 12 science and 12 social studies textbooks adopted by school districts around California. As each textbook varies in the kind of pedagogically complementary sections it contains (e.g. review, assessment, enrichment), only the main texts were chosen for this analysis. The textbooks were published by 5 different publishers (3 of which contributed both science and social studies textbooks), with 3 grade levels per publisher (6^{th}/7^{th}/8^{th} grade).

The corpus in its entirety consists of 1,667,906 words across 125,831 sentences (55,005 sentences for science, 70,826 sentences for social studies). We treated sentence boundaries as an imperfect proxy for proposition boundaries, using punctuation to automatically separate sentences (period, question mark, or exclamation point). Though this approach may both underestimate the number of unique propositions (when a single sentence conjoins multiple clauses) or overestimate that number (when a punctuation mark serves a purpose other than clause division, e.g., the period in "Ms."), we have no reason to believe that the inaccuracies will be more prevalent in one subject area compared to another.

Methods

In order to test what factors influence whether or not a discourse marker will be present and what type of marker will be used, we conducted two separate analyses. The first models the binary outcome of DM presence/absence in each sentence, testing how that outcome is influenced by two fixed factors and their interaction: subject area (Science/SS) and grade level (6th/7th/8th). The second models the proportions of different marker types, testing how that proportion is influenced by three fixed factors and their interactions: marker category (contrastive/elaborative/inferential/temporal), subject (Science/SS), and grade (6th/7th/8th).

Both models treat textbook section (as opposed to unit or chapter) as the unit of repeated measure, so that each section's sentences are associated with a grouping ID representing the information of publisher/subject/unit/chapter/section. The first model predicts the raw DM presence/absence outcome in a mixed effects logistic regression with the fixed factors listed above as well as random effects for publisher and unique ID nested within publisher. For the second model, the raw data is collapsed by unique ID to form proportions (e.g., "what proportion of sentences in a particular textbook section were elaborative?"), and the elogit proportions are analyzed in a mixed effect linear regression (http://talklab.psy.gla.ac.uk/tvw/elogit-wt.html). Random by-publisher slopes for grade and DM category were included where possible. Prior to analysis, fixed factors were centered to reduce collinearity and to enhance the interpretability of estimates of coefficients. All models were fit using the lmer function in the lmer4 package in R (Baayen et al., 2008; lmer4 version 0.999375-42), using maximum-likelihood (ML) estimation. For each factor and interaction in the logistic regression, we report the coefficient estimate and p-value (based on the Wald Z statistic; Agresti,2002). For the linear regression, we report the coefficient estimate, the t-value, and a p-value obtained using a model comparison approach, based on a likelihood-ratio X2 (df=1) test of the change in the goodness of fit between the full model and a comparison model in which only the relevant fixed effect or interaction was removed.

Results

For the primary question of whether subject area influences DM presence/absence, the logistic regression did indeed show that Science texts yielded a higher rate of DM usage (16.2% of sentences contained a DM) than SS texts (14.1%; Coeff=-0.175; $p<0.001$). The effect of grade level was marginal with DM usage increasing slightly across 6th grade (14.4%), 7th grade (14.6%), and 8th grade (15.8%; Coeff=0.063; $p=0.10$). There was also a subject area X grade level interaction whereby science showed a greater increase across grade levels than SS (Coeff=-0.179; $p<0.001$). The pattern of results is shown in Table 1, broken down by subject area and grade level.

	Science	**Social Studies**
6th grade	0.149	0.140
7th grade	0.147	0.145
8th grade	0.190	0.138

Table 1: Rates of DM usage per sentence (0=absent; 1=DM present)

Modeling the binary outcome of DM presence/absence does not permit us to test how the rate of different DM categories varies across subject area and grade level (because the question of how DM category influences DM usage is not defined for the sentences with no DM). To address this, we calculated the proportion of sentences in each section (with its unique ID) that contained a contrastive marker, an elaborative marker, an inferential marker, and a temporal marker. The means of these by-section proportions are shown in Figures 1 and 2 to illustrate how the distribution of DM categories varies by subject area and grade level. Note that, in each grade level for each subject area, the 4 colored bars (i.e., the 4 DM categories) sum together to give the overall rate of DM usage that is depicted in the corresponding cell of Table 1.

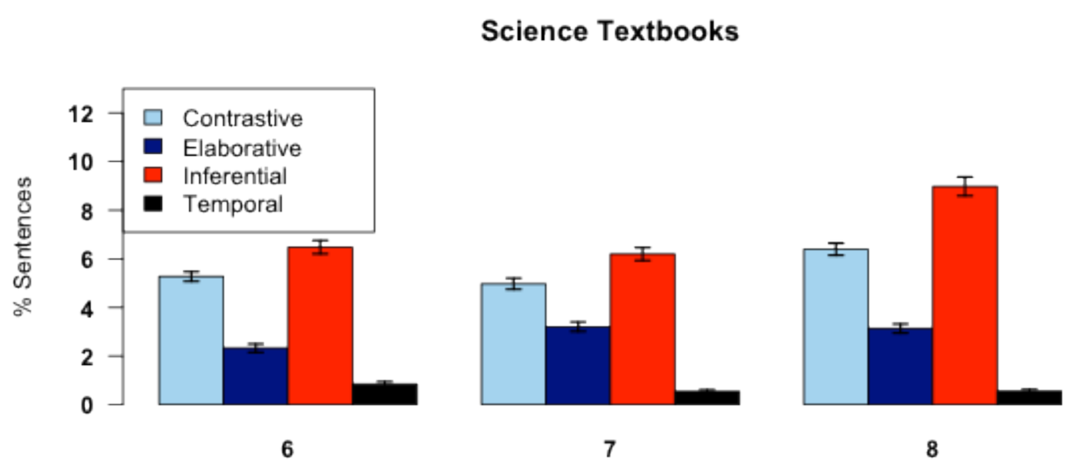

Fig 1: Science rate of DM usage across DM categories, calculated by section (0% is no sentences in a section; 100% is all sentences in a section)

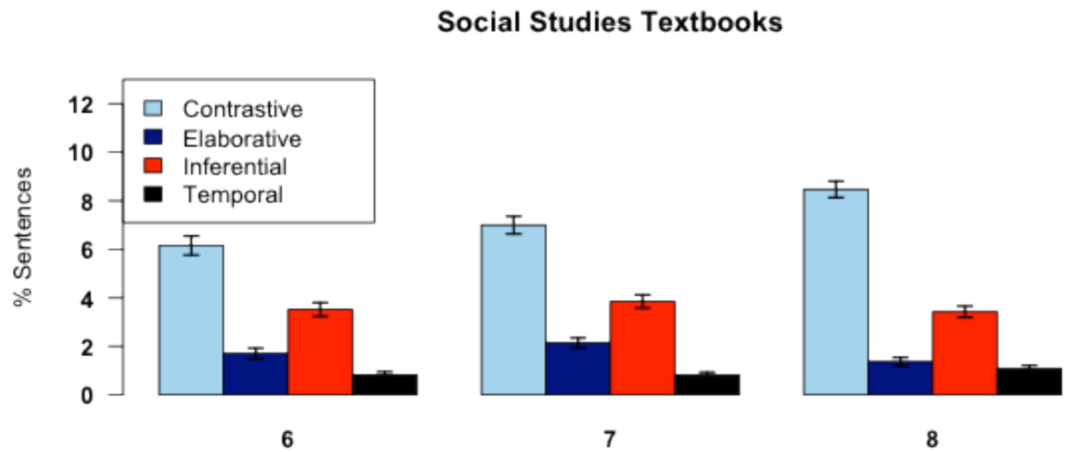

Fig 2: Social Studies rate of DM usage across DM categories, calculated by section (0% is no sentences in a section; 100% is all sentences in a section)

For the question of what factors influence the proportion of DMs that are present, the linear regression again showed a main effect of subject area, whereby science yields slightly higher proportions of DMs than social studies (Coeff= -0.123; t-val=-11.70; p<0.001). The slight increase in the proportion of DMs across grade levels reaches significance in this model (Coeff=-0.028 ; t-val=-5.14; p<0.05). As Figures 1 and 2 show, not all DM categories are used with equal frequency. Treating Elaborative markers as the baseline, the model shows that there are significantly more Contrastive (Coeff=0.883; t-val= 36.86; p<0.001) and Inferential (Coeff=0.404; t-val= 13.12; p<0.001) markers than Elaborative and significantly fewer Temporal markers (Coeff=-0.878; t-val=- 21.85; p<0.001).

As Figures 1 and 2 also show, the most striking difference between the subject areas is the different rates of usage of the 4 DM categories. This is apparent in the significant subject-area X marker-category interactions: compared to the Elaborative baseline, Contrastive markers are more frequent in SS than in

Science (Coeff=0.429; t-val=32.89; p<0.001) as are Temporal markers (Coeff=0.453; t-val=25.70; p<0.001), whereas Inferential markers are more frequent in Science than in SS (Coeff=-0.506; t-val=-30.05; p<0.001). There are also differences by grade level. For one, proportions increase across grade levels for Science but not SS (subject-area X grade-level interaction; Coeff=-.118; t-val=-14.01; p<0.001). Additionally, the proportion of Contrastive markers increases over grade levels (Coeff=0.083; t-val=15.24; p<0.001), whereas Temporals decrease (Coeff=-0.046; t-val=-6.98; p<0.001); the proportion of Inferential markers doesn't change reliably over grades when the data is collapsed across subject area (Coeff=0.007; t-val=1.08; p>.10), but the subject-area X grade-level X marker-category interactions help clarify this. The Contrastive-marker increase over grade levels is limited to SS (subject-area X grade-level X Contrastive interaction: Coeff=0.106; t-val=9.24; p<0.01), whereas the Inferential-marker increase over grade levels is limited to Science (subject-area X grade-level X Contrastive interaction: Coeff=0.117; t-val=- 8.46; p<0.01), as is the Temporal-marker decrease (subject-area X grade-level X Temporal interaction: Coeff=0.164 ; t-val=11.64; p<0.001).

Scholarly Significance

Among the literacy skills students require to access science concepts, Yore and Shymansky (1991) state that the ability to “read about science is a critical skill to have in order to develop scientific literacy” (p.29). Furthermore, as almost all of what we call “knowledge” is based on language (Wellington&Osborne,2001; Hines,Wible,&McCartney,2010) and that, in the end, *doing* science depends on being able to *talk* science to ourselves and to others (Lemke,1990). Our findings about discourse-marker usage offers the education community crucial information about the variation and challenges inherent to the language which students encounter in textbook in different content areas.

Much research still needs to be done to determine how addressing language factors in science instruction at the secondary level can foster science achievement (Greenleaf et.al.,2011; Fang&Wei,2010). However, understanding features of the text itself could be used to refine professional developments for science teachers so they can tailor instruction appropriately for students' literacy and comprehension. For example, the significantly higher frequency of inferential markers and lower frequency of contrastive markers might mean that science texts present science mostly as a set of results rather than as processes in which differing opinions had to be reconciled. Thus, science teachers would need to supplement the textbook with instruction and materials that model the crucial role that argumentation plays in scientific findings (Wellington&Osborne,2001). Finally, with an understanding of the language of content-area textbooks, teachers can implement specific discourse practices that better support students' comprehension.

References

Agresti, A. (2002). Categorical data analysis (2nd ed.). Wiley.

Asher, N., & Lascarides, A. (2003). Logics of Conversation. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language (Special issue on Emerging Data Analysis), 59, 390–412.

Barr, D. (2010). "Walkthrough of an 'empirical logit' analysis in R."

Online: http://talklab.psy.gla.ac.uk/tvw/elogit-wt.html.

Fang, Z. and Wei, Y. (2010). "Improving Middle school science literacy through reading infusion." *The Journal of Education and Research. 103*, 262-273.

Fraser, B. (2006). "Towards a theory of discourse markers." *Approaches to Discourse Particles.* Ed. Kerstin Fischer. New York: Elsevier. 189-204. Print.

Glencoe (2007). *Focus on physical science.* New York: Glencoe/McGraw Hill.

Greenleaf, C., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., Schneider, S. A., Madden, S., & Jones, B. (2011). Integrating Literacy and Science in Biology: Teaching and Learning Impacts of Reading Apprenticeship Professional Development. *American Educational Research Journal, 48*(3), 647-717.

Grimes, J. (1975). The thread of discourse. The Hague: Mouton.

Halliday, M. & Hasan, R. (1976). *Cohesion in English.* New York: Longman.

Hines, P., Wible, B., McCartney, M. (2010). Learning to read, reading to learn. *Science, 23*(328), 447.

Hobbs, J. R. (1979). Coherence and coreference. Cognitive Science, 3, 67–90.

Jetton, T. & Shanahan, C. (2012). Preface. In T. Jetton & C. Shanahan (Eds.) *Adolescent literacy in the academic disciplines* (pp. ix-xii). New York: The Guilford Press.

Kehler, A. (2002). Coherence, reference, and the theory of grammar. Stanford, CA: CSLI Publications.

Lemke, J. (1990). *Talking science: language learning, and values*. CT: Ablex.

Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. Text, 8, 243-281.

McNamara, D. (2001). Reading both high and low coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology, 55,* 51-62.

Nemo, F. (2007). Reconsidering the discourse marking hypothesis. Even, even though, even if, etc. as morpheme/construction pairs. In A. Celle & R. Huart (Eds.) *Connectives as discourse landmarks* (pp. 195-210). Philadelphia: John Benjamins Publishing Company

Polanyi, L. (1988). A formal model of the structure of discourse. Journal of Pragmatics, 12, 601-638.

Shanahan, T. & Shanahan, C. (2008). Teaching Disciplinary Literacy to Adolescents: Rethinking Content-Area Literacy. *Harvard Educational Review*. *78(1), 40-59.*

Taboada, M. & Mann, W. (2006). Rhetorical structure theory: looking back and moving ahead. *Discourse Studies.* 8(3), 423-459.

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J.

(2003). Looking inside the classroom: A study of K-12 mathematics and science education in the United States. *Horizon Research Inc., Chapel Hill, NC, retrieved from http: www.horizonresearch.com*.

Wellington, J. & Osborne, J. (2001). *Language and literacy in science education*. England: Open University Press.

Yore, L., & Shymansky, J. (1991). Reading in science: developing an operational conception to guide instruction. *Journal of Science Teacher Education, 2*(2), 29-36.

Appendix I

Contrastive Markers

but, alternatively, although, conversely, despite (this/that), even so, however, in spite of, in comparison, in contrast, instead, nevertheless, nonetheless, notwithstanding, on the other hand, on the contrary, rather, regardless, still, though, whereas, yet, similarly

Elaborative Markers

above all, alternatively, analogously, besides, correspondingly, equally, for example, for instance, further, furthermore, in addition, in other words, in particular, likewise, more accurately, more importantly, more precisely, moreover, on that basis, otherwise, rather, similarly

Inferential Markers

so, after all, all things considered, as a conclusion, as a consequence of, as a result, because of, consequently, for this reason, for that reason, hence, it follows that, in this/that/any case, on this/that condition, on these/those grounds, then, therefore, thus, when you, equally

Temporal Markers

eventually, finally, immediately, afterwards, in the meantime, meanwhile, originally, subsequently, lastly