Discourse Processes

ISSN: 0163-853X (Print) 1532-6950 (Online) Journal homepage: https://www.tandfonline.com/loi/hdsp20

Coherence Relations in Primary School Textbooks: Variation across School Subjects

Romualdo Ibáñez, Fernando Moncada & Benjamín Cárcamo

To cite this article: Romualdo Ibáñez, Fernando Moncada & Benjamín Cárcamo (2019) Coherence Relations in Primary School Textbooks: Variation across School Subjects, Discourse Processes, 56:8, 764-785, DOI: 10.1080/0163853X.2019.1565278

To link to this article: https://doi.org/10.1080/0163853X.2019.1565278

Published online: 23 Jan 2019.

DISCOURSE PROCESSES
2019, VOL. 56, NO. 8, 764–785
https://doi.org/10.1080/0163853X.2019.1565278

Coherence Relations in Primary School Textbooks: Variation across School Subjects

Romualdo Ibáñez[a], Fernando Moncada[b], and Benjamín Cárcamo [b]

[a]Instituto de Literatura y Ciencias del Lenguaje Pontificia Universidad Católica de Valparaíso, Viña del Mar, Chile; [b]Instituto de Literatura y Ciencias del Lenguaje, Pontificia Universidad Católica de Valparaíso, Viña del Mar, Chile

ABSTRACT

The aim of this study was to compare the use of coherence relations in the school textbooks of four primary school subjects. A corpus of 1,882 texts, corresponding to the genres used to teach in the school textbooks of Science, Language, Mathematics, and History, was analyzed. The analysis was performed manually using a taxonomy that incorporates and adapts criteria developed in the Cognitive approach to Coherence Relations and in the Connectivity Model. Results show that while some relations are used across school subjects (Conjunction, Concept Description), others are used almost exclusively in certain school subjects (Condition-Event in Science, Basic Contrast in History, Deictic in Language, and Condition-Question in Mathematics). Considering that types of coherence relations are different in terms of internal complexity, these results may represent relevant information for classroom instruction, primary school textbook design, and text comprehension research.

Introduction

Given that most learning in school is done through written language, school textbooks (STs) play an important role in education (Daneman, 1991). Particularly in Chile, teachers recognize it as the most important tool during the teaching learning process (MINEDUC, 2012). Unfortunately, studies conducted in different countries have shown that students do not comprehend much when they read from their STs (Graesser, McNamara, & Louwerse, 2003; Knoepke et al., 2017; van Silfhout, Evers-Vermeul, & Sanders, 2015). Researchers have suggested different reasons for students´ low comprehension levels: lack of linguistic cues in the textbook for students to create coherent representations (Otero, León, & Graesser, 2002), lack of match between the students´ reading skills and the material assigned (Crossley, Greenfield, & McNamara, 2008; Dufty, Graesser, Louwerse, & McNamara, 2006), lack of previous knowledge about the domain (McNamara, 2001, McNamara, Kintsch, Songer, & Kintsch, 1996), and so on. What all these possible reasons have in common is that to comprehend and learn from STs, students need to establish coherence relations, that is, the meaning relations that connect two discourse segments (minimally clauses) (Sanders & Noordman, 2000). Indeed, different studies have proved that coherence relations represent the building blocks of discourse comprehension and learning (Graesser et al., 2003; Louwerse, 2001; Sanders, Spooren, & Noordman, 1992).

Hence, it seems that to improve comprehension and learning from STs, students should be taught how to process coherence relations. However, before starting that enterprise it is necessary to

CONTACT Romualdo Ibáñez romualdo.ibanez@pucv.cl Instituto de Literatura y Ciencias del Lenguaje, Pontificia Universidad Católica de Valparaíso, Av. El Bosque 1290, 5º piso, oficina 5-17, Campus María Teresa Brown de Ariztía, Sausalito, Viña del Mar, Chile

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hdsp.

understand that not all types of coherence relations in STs will impose the same degree of difficulty to learners. Psycholinguistic research has proved that there are types of coherence relations that are processed faster than others (Canestrelli, Mak, & Sanders, 2013; Traxler, Bybee, & Pickering, 1997; Zufferey & Gygax, 2016). Similarly, it has been reported that there are types of coherence relations, such as causal (Peter arrived late. He had missed the bus.), which are easily remembered (Black & Bower, 1980; Graesser & Clark, 1985; Trabasso & Sperry, 1985, van Den Broek, 1990). It has also been shown that explicit relations (Peter arrived late *because* he missed the bus) involve a lower cognitive load than implicit relations (Peter arrived late. He had missed the bus), enabling a deeper and longer lasting comprehension (Maury & Teisserenc, 2005; Millis & Just, 1994; Sanders & Noordman, 2000). All these differences could be explained by the cumulative cognitive complexity approach (Evers-Vermeul & Sanders, 2009; Spooren & Sanders, 2008). This approach suggests that coherence relations differ in internal complexity, which not only affects processing but also acquisition. Causal relations, for example, are assumed to be more difficult to process than additive ones, while positive relations are assumed to be easier to process than negative ones. Regarding acquisition, less complex relations, such as positive-causal coherence relations, are assumed to be acquired earlier than negative-causal coherence ones (Bloom, Lahey, Hood, Lifter, & Fiess, 1980; Evers-Vermeul & Sanders, 2009).

In this scenario some questions come up: Which are those coherence relations that students will more frequently find in STs? How are coherence relations instantiated in STs? Is knowledge communicated through particular coherence relations in STs? Before being tempted to provide an answer, it is necessary to consider that coherence relations use varies due to different causes, such as text type, language, discourse genre, and so on. Hence, it is also necessary to consider that the type of coherence relations students have to face in STs may vary depending on school subject and level. This assumption is based on corpus studies that have shown this type of language variation. Sanders (1997), for example, noted that there are significant differences between informative and non-informative texts in terms of the most frequently occurring coherence relations. Pretorius (1994) showed that narrative texts and scientific texts differ in the dominant types of causal coherence relations. Similarly, Taboada and Gómez-González (2012) observed that the dominant rhetorical function of a given type of coherence relation varies between oral and written texts. Also, Redeker and van der Vliet (2014) observed variation in terms of frequency between expository and persuasive genres. In Spanish, Ibáñez, Moncada, and Santana (2015) noted disciplinary variation in in the frequency of coherence relations used in academic genres of Biology and Law. Ibáñez and Moncada (2017) also found disciplinary differences in the frequency of coherence relations used in academic abstracts written in Spanish.

Against this background the current study compares the use of coherence relations in the STs of four primary school subjects. We analyzed a corpus composed of 1,882 texts, corresponding to the genres used to communicate knowledge (Knowledge Genres [KGs]) in the STs of Mathematics, Language, History, and Science. These school subjects are declared to be the core subject areas in the National Curricular Bases proposed by the Chilean Ministry of Education (MINEDUC, 2012). Additionally, the motivation for analyzing STs of primary school is that different studies (Cain, Patson, & Andrews, 2005; Knoepke et al., 2017) suggest that primary school students still struggle to process certain types of coherence relations. The corpus was analyzed manually using a taxonomy that integrates and adapts criteria proposed by Sanders et al. (1992) along with some criteria described in Renkema´s Connectivity Model (2009).

The article is organized as follows. First, we provide an introduction to the concept of ST and KG. Second, we first describe the criteria proposed in the models we based our taxonomy and, then, our approach to coherence relations. Third, we present the methods. Fourth, we present and discuss the results, including the statistical distributions of the coherence relations across school subjects. Finally, we discuss the implications of those results and provides the conclusion.

ST and KGs

In recent years the study of STs has advanced considerably. Scholars from areas such as language and education have studied their contents (Atienza, 2007; Ramírez, 2012), teaching methods (Barrios, 2012; Marzábal, 2012; Sáiz, 2011), publishing market (Bourdieu, 1999; Petrucci, 2001), and, more recently, discursive configuration (Barletta & Chamorro, 2011; Chamorro, Mizuno, & Moss, 2011). From a discursive point of view, STs are understood as discourse macro genres, whose main communicative purpose is to instruct declarative and procedural knowledge related to a particular subject matter (Martin & Rose, 2008; Rose, 2014). This discourse genre-oriented type of research has shown that STs are composed of two sets of genres, through which the macro genre's communicative purpose is fulfilled: the KG, through which academic knowledge is presented (Definition, Dictionary, Map, etc.), and Curriculum Genres, through which KGs are acquired (Syllabus, Lesson Plan, etc.) (Christie, 2002). According to Martin and Rose (2013), each discipline uses different KGs to communicate knowledge, which can be observed in contexts of schooling and, more precisely, in STs. To describe this type of variation in STs, Ibáñez, Moncada, Cornejo, and Arriaza (2017) analyzed a corpus consisting of 100 STs from Mathematics, Language, History, Science, and English. Fifiteen KGs were identified (*Definition, Biography, Note, Procedural Guide, Content Exposition*, etc.), which allowed the classification of 11,139 instances. Another interesting finding of that study was that the predominant genre was different in each subject: In History, the dominant one was *History Source*; in Mathematics, *Procedural Guide*; in Language, *Definition*; and in Science, *Content Exposition*. It was also observed that there are genres that occur exclusively in one subject: *History Source* occurs only in History, *Scientific News* exclusively in Sciences, and *Frequent Expressions* only in English STs. It was also observed that although a large part of the KGs occurred across school subjects, their frequency of occurrence varied across school subjects. For example, the genre *Procedural Guide* was found across the school subjects but half of its occurrences were concentrated in Mathematics. This was interpreted by the authors as one indicator of variation in the way in which school subjects are configured at a discourse level.

Coherence relations and their classification

Given its relevance to communication in general and to discourse studies in particular, coherence relations have been a topic of interest for scholars from different disciplines over the last four decades. This interest has given rise to discussions regarding their nature (Bublitz, 1998; Redeker, 2000; Sanders & Pander Maat, 2006), their definite number (Grosz & Sidner, 1986; Hovy, Lavid, Maier, Mittal, & Paris, 1992; Mann & Thompson, 1988; Sanders et al., 1992), annotation methods (Hoek, Evers-Vermeul, & Sanders, 2017; Scholman, Evers-Vermeul, & Sanders, 2016; Taboada & Das, 2013), time and order of acquisition (Bloom et al., 1980; Evers-Vermeul, 2005; Evers-Vermeul & Sanders, 2009), and so forth. Such diversity in views has led to a large number of theoretical models and taxonomies, which have been divergent in some cases and, in others, complementary (Louwerse, 2001). In the current study we used a taxonomy based mainly on the theory proposed by Sanders et al. (1992) complemented by categories proposed by Renkema (2009). Sanders et al.´s taxonomy (1992) was selected because it provides researchers with four cognitive primitives that allow the classification of multiple coherence relations. It is worth mentioning that this proposal has been supported by strong empirical evidence (Sanders & Noordman, 2000; Sanders & Pander Maat, 2006). On the other hand, Renkema's proposal was chosen on the grounds that it accounts for further specification of the classification of additive relations, which in Sanders et al.´s (1992) are not fully developed.

Sanders et al. (1992) claimed that every single coherence relation can be classified by integrating four cognitive primitives: Basic Operation, Source of Coherence, Order of Segment, and Polarity. Basic Operation indicates the strength of the semantic link between the discourse segments. The strongest links are *Causal* (because they are connected by an implication operation), whereas the

weak ones are *Additive* (because they are connected in a logical conjunction). The Source of Coherence is related to the nature of the link that is established. In the original proposal the authors identified two types of Source of Coherence: *Semantic*, when the link is established at the level of propositional content, and *Pragmatic*, when it is established at the level of illocutionary meaning. In more recent proposals (Spooren & Sanders, 2008) this classification has been reformulated. *Semantic* relations have been understood as Content relation, whereas *Pragmatic* relations were subdivided into *Speech Acts* (links motivated by illocutionary force) and *Epistemic* ones (connections that involve logical reasoning and inferences). Order of Segments accounts for the correspondence between each discourse segment and its role as either antecedent (P) or consequent (Q) in the coherence relation. Thus, there is *Basic Order* when P corresponds to the first segment (S1) and Q to the second one (S2). Conversely, the order is *Nonbasic* when the opposite sequence is present. Finally, Polarity distinguishes between *Negative* and *Positive* relations. A negative relation includes a negation of the propositional content of one of the segments while a positive relation does not. By integrating the four cognitive primitives, the relation existing in "Several roof tiles fell from the roof, but I still think that the storm was not that strong" could be characterized as causal, epistemic, nonbasic, and negative, thus allowing the classification of this relation as Concession (Spooren & Sanders, 2008).

The application of these four primitives allows the successful classification of a variety of relations. The classificatory capacity and psychological plausibility of this model has turned it into a widely used tool in empirical research (Evers-Vermeul & Sanders, 2011; Scholman et al., 2016; Spooren & Sanders, 2008; van Silfhout et al., 2015). However, this proposal shows limitations when discriminating between cases such as (1) and (2) below

(1) [Juan bought a computer.] [I bought a phone.]
(2) [Juan bought a computer.] [It has a decent processing power.]

According to Sanders et al. (1992), the coherence relations holding in (1) and (2) can be described as *Additive, Content*, and *Positive*, and, in turn, classified as *Joint*. Despite this, it can be noticed that the nature of the relation in (1) and (2) is different. In (1) segments are linked because they only share the topic, whereas in (2) the relation can be established because the information expressed in the second segment complements what is expressed in the first one. To classify cases such as (2), some categories from the level of Adjunction of the Connectivity Model (Renkema, 2009) have been proved to be useful (Ibáñez et al., 2015).

Renkema (2009) develops a model that describes how discourse is connected at three levels: Conjunction, Adjunction, and Interjunction. At the level of Adjunction, discourse is connected by adding information, where one proposition adds information to another. In the model a proposition is understood as composed, at least, of a concept (a person or thing), generally realized by a noun phrase, and an event (an action or process), realized by a verb. At the level of Adjunction, propositions can be connected in three different ways, depending on the component to which information is added: the concept (C), the Concept plus the Event (C + E), or the whole Event (E). When information is added to a Concept, the procedure is called Elaboration (*This year there were presidential elections. This is the most important civic instance for a citizen*); when it is added to a Concept plus the Event, it is called Extension (*This year there were presidential elections. Results show that people were not happy with the former president*); and, when it is added to the Event, it is called Reinforcement (*This year there were presidential elections. They took place in November*). These three types of procedures allow the classification of three families, eight categories, 20 types, and 16 subtypes of relations. Relation in (2) is an exemplar of the Elaboration category, as the second discourse segment (*it has a decent processing power*) provides information about a concept in the first (*computer*).

In the current study we used a taxonomy based on the one developed by Sanders et al. (1992, 1993), complemented with modified categories from the Adjunction level of Renkema's Connectivity

Model (2009). As we understand that Relational and Additive coherence are different in nature, we propose different criteria for their description. For describing Relational Coherence, we used a modified version of the model proposed by Sanders et al. (1992, 1993). These minor modifications were applied to some of the primitives. In the case of Basic Operation, we distinguish between Adjacency relations (when two events only share a topic) and Implication relations (when, in addition to sharing the topic, there is an implication between the events). Content relations were divided into three types: Volitional (when Q is mediated by an intention), Deontic (when Q corresponds to a mandatory action), and Neutral (when Q is an unintended action or state). As a result of the integration of these criteria and their indicators, a specific type of coherence relation can be identified and classified (see Appendix A).

For describing Additive Coherence, we adopted the notion of information complementation used at the level of Adjunction of Renkema's model. The categories of our taxonomy are slightly different because they resulted from the integration of two criteria: Information Focus and Information Type. The former refers to the element to which information is added, which can be a Concept (C) (a person or thing) of an event, the Action (A) (an act, state or process) of an event, or the whole Event (C + A). The latter corresponds to the type of information that is being provided, which can be Adjectival (when it corresponds to some characteristic of a Concept or Event), Adverbial (when it corresponds to the information about how, when or where of the Action of an event), or Metadiscursive (when the Event that adds information refers to the structure or content of the text). The application of these criteria and indicators leads to the identification of four categories of additive relations: Elaboration, Extension, Circumstantial, and Metadiscursive. Each of these categories groups several types of relations (see Appendix A).

By applying the taxonomy used in the current study, example (1) is identified as a case of Relational Coherence. Hence, through the application of the criteria for Relational Coherence, the Basic Operation is Adjacency (since there is no implication between events), the Source of Coherence is Content Neutral, and the Polarity is Positive. Altogether, these criteria configure a Conjunction relation. The coherence relation in (2), on the other hand, is an example of Additive Coherence as the second event (*it has a decent processing power*) provides information about an element of the previous event. Therefore, the Information Focus is the Concept (*computer*) and the Information Type is Adjectival as it corresponds to a concept attribute. Thus, example (2) corresponds to the category of Elaboration and to the type Concept Description.

Considering variation as a phenomenon inherent to discourse, the hypothesis of this study was that coherence relations in STs vary according to the school subject. More precisely, we expect that those types of coherence relations that are frequent in a school subject will not necessarily be frequent in another one. Based on this assumption, the aim of this study was to compare the use of coherence relations in the STs of four school subjects. A second goal was to determine if there are coherence relations that could be regarded as prototypical of a particular school subject.

Methods

Corpus

The corpus consisted of 1,882 texts corresponding to the two most frequent KGs per school subject, previously identified by Ibáñez et al. (2017) in STs of Language, Mathematics, Science, and History. The STs examined by Ibáñez et al. (2017) were published by leading publishing companies in Chile (Cal y Canto, Galileo, Santillana, Zig Zag, Piedra de Sol, SM) and distributed by the Chilean Ministry of Education to the primary school students, who attended public and subsidized-by-the-government schools during 2012 and 2016 (see Appendix B for a sample of the STs from which the KGs were identified). Primary school in Chile goes from first to eighth grade and students should finish it at the age of thirteen.

Table 1 presents the definition of the KGs analyzed in the current study (see Appendix C for an example of a Procedural Guide in Mathematics). In the construction of the corpus, the sample size

Table 1. Definition of the KGs Analyzed in the Study.

KG	Definition
Definition (DEF)	Discourse genre whose communicative purpose is to present the meaning of a particular concept. Its discourse organization is expository and its semiotic organization is verbal.
Content exposition (CE)	Pedagogical genre whose communicative purpose is to develop key topics specific of a subject matter. Its discourse organization is expository or narrative and its semiotic organization is verbal and graphic.
Procedural guide (PG)	Pedagogical genre whose communicative purpose is to instruct the execution of procedures specific of a subject matter. Its discourse organization is instructional and its semiotic organization is verbal and graphic.
Note (NOT)	Pedagogical genre whose communicative purpose is to present specific information about a topic. Its discourse organization is expository and its semiotic organization is verbal and graphic.
History source (HS)	Pedagogical genre whose communicative purpose is to organize contents based on milestones related to a phenomenon. Its discourse organization is expository and its semiotic organization is verbal and graphic.

was calculated with a margin of error of 5% based on the total number of identified texts per genre in each school subject textbook. Texts were randomly selected and their distribution is detailed in Table 2.

Framework of analysis

Under the assumption that relation inventories are open (Das & Taboada, 2017; Taboada & Das, 2013) and also considering that the aim of our study involved contrastive interests, we adopted a top-down bottom-up framework (Biber, Connor, & Upton, 2007), which allowed us to obtain an updated taxonomy, *ad hoc* to the corpus of analysis. The framework was composed of two stages. In the top-down stage we carried out a literature review to select appropriate criteria (and taxonomy) for analyzing the corpus. We initially selected the criteria proposed by Sanders et al. (1992). In the bottom-up stage we applied these preliminary criteria to a sample corpus (15%), finding that they were not sensitive enough to account for the diversity of relations identified (see examples (1) and (2)). Therefore, it was necessary to go back into the top-down stage again and look for complementary criteria that allowed us to classify those relations that could not be properly classified by the initial criteria (and taxonomy). After a second literature review, the complementary criteria were found in Renkema's Connectivity Model (2009). As a result, the taxonomy was updated and then validated (see Validation below). After that, the analysis was performed in the whole corpus. This analysis was carried out by three pairs of coders (see Manual annotation); all undergraduate and graduate students in the field of discourse studies, who were instructed over a period of four months in how to conduct the annotation process.

Manual annotation

To illustrate the process, two examples extracted from the corpus are presented below. For demonstrative purposes, the extracts have been segmented into their two constitutive events (E).

Table 2. Distribution of knowledge genres in different subjects.

Subject	Knowledge Genre	Total Frequency
Language	Definition (DEF)	157
	Content exposition (CE)	118
Mathematics	Procedural guide (PG)	191
	Note (NOT)	109
History	History source (HS)	702
	Content exposition (CE)	319
Science	Content exposition (CE)	177
	Note (NOT)	109
Total		1,882

(3) $_{E1}$ Se rompen los enlaces entre las partículas de los materiales. Por lo tanto, $_{E2}$ estas se separan y experimentan nuevas combinaciones. (Ciencias)

$_{E1}$ *The bonds between the particles of matter are broken. Therefore,* $_{E2}$ *they are separated and experience new combinations.* (Science)

(4) $_{E1}$ La Zona Central se extiende desde el río Aconcagua hasta el río Biobío. $_{E2}$ Esta área se caracteriza, entre otros aspectos, por sus marcadas estaciones, fértiles suelos. (Historia)

$_{E1}$ *Central Zone extends from Aconcagua River to Biobío River.* $_{E2}$ *This area is characterized, among other aspects, because of its marked seasons and fertile soils.* (History)

The first step is to identify the type of coherence (Relational or Additive). In (3) a case of Relational Coherence holds because E1 and E2 share the same topic in the first place. Therefore, the criteria for this type of relation are applied. The Basic Operation is Implication because the fact described in E1 constitutes the antecedent (P) of the fact presented in E2, which constitutes the consequent (Q). The Source of Coherence is Content Neutral because the consequent (Q) corresponds to a state-of-affairs in an unintended ideational plane. The Polarity is Positive, given that Q corresponds to a state-of-affairs that is consistent with what could be expected on the basis of P. Finally, the Order of the Events is Basic because E1 constitutes the antecedent and E2, the consequent. The combination of these criteria and indicators configures the Cause-Effect relation (see Appendix A).

In (4), E2 provides additional information for (part of) what is described in E1. Thus, the criteria for additive coherence are applied (see above). Regarding the Information Focus, E2 provides further information related to a concept in E1 (Central Zone). The Type of Information is Adjectival since such information is about the characteristics of the concept. The combination of these criteria and indicators configures the Concept Description relation (see Appendix A).

Validation

To determine construct validity, the new categories and types of coherence relations were validated by six experts (four students and two professors from a Linguistics graduate program). They were asked to evaluate each category in terms of label, definition and example (extracted from the corpus). Their evaluations were analyzed based on the coefficient of agreement for nominal scales (Cohen, 1960). A minimum index of 90% of agreement was achieved for each aspect.

Interrater agreement

For the sake of ensuring interrater agreement, two complementary strategies suggested by Spooren and Degand (2010) were followed. First, in the two-coders-discuss strategy, each fragment was analyzed independently by the two coders in each couple, and afterwards they discussed the differences. In the case of disagreement, the differences were discussed until agreement was reached. Subsequently, following the same principle, the results of the analysis of each pair was evaluated by another pair of coders.

The second strategy used to confirm interrater agreement was estimating Fleiss' kappa coefficient. We decided to use this coefficient because it permits the estimation of agreement between more than two raters (Fleiss, 1981), which was the case in this study. The agreement index in this study was $\kappa = 0.68$. Landis and Koch (1977) stated that an index of this magnitude (0.61–0.80) can be interpreted as substantial. In addition, for coherence relations studies, Spooren and Degand (2010) have claimed that due to the level of subjectivity involved in this kind of analysis, kappas tend to be lower. The acceptance of this phenomenon has been termed as the expectation of "soft kappas."

Data analysis and processing

After annotating the corpus, descriptive and inferential statistics were used for the analysis. The frequency distributions for Adjacency, Implication, and Additive relations were calculated first. Once this was done, statistically significant differences between school subjects and the type of coherence relations were made using a χ^2 test and the Bonferroni correction for the performance of the post-hoc tests (MacDonald & Gardner, 2000). Finally, prototypical coherence relations were determined for each of the subjects.

Results and discussion

Here we first describe the types and frequency of the coherence relations identified in the total corpus and then present their distribution across school subjects.

Types of coherence and coherence relations

The total number of coherence relations identified was 6,677. Table 3 illustrates their distribution per type of coherence (Relational vs. Additive) and category.

Table 3 shows that the highest amount of coherence relations is concentrated in the categories of Implication and Adjacency, which together make up Relational Coherence (62.93%). This figure greatly exceeds the occurrence of the other four categories, which are part of Additive Coherence (37.07%). The three highest frequencies in the six categories are found in Implication (33.91%), Adjacency (29.02%), and Elaboration (21.36%). On the contrary, Metadiscursive and Circumstantial coherence relations show the lowest frequencies. For the six categories 40 types of coherence relations were found (see Appendix D). Table 4 shows the types whose frequency is higher than 1%.

Table 4 shows that 6 types of relations concentrated 53% of the coherence relations found, whereas the remaining is distributed among 34 types (14 of which have a frequency less than 1%). These results can be taken to indicate that some types of coherence relations play a central role in the dissemination of knowledge in STs, whereas others are entirely peripheral. Examples of the former are Nonbasic Contrast (0.3%), Action-Reason (0.4%), and Effect-Cause (0.9%), among others (see Appendix D for the complete list).

As stated earlier, the main communicative purpose of STs (to instruct declarative and procedural knowledge related to a subject matter) is fulfilled mainly by its constitutive KGs. Therefore, it would be assumed that the descriptions, explanations, definitions, classifications, and so on exposed in those KGs are presented through types of coherence relations belonging to different categories. This is observed in Table 4, where among the most frequent types of coherence relations, two belong to the category of Adjacency (Conjunction, 19.48%; Sequence, 6.86%), three to Implication (Cause-Effect, 5.3%; Condition-Event, 4.06%; Claim-Argument, 4.03%), and two to Elaboration (Concept Description, 12.95%; Concept Specification, 4.57%). The predominance of Conjunction is consistent with the results reported in different corpus-based studies which have shown that this relation is among the most frequent in academic genres (Ibáñez & Moncada, 2017). A possible reason for the high frequency of Conjunction in STs could be related to the fact that, as observed in some

Table 3. Distribution of Coherence Relations by Type of Coherence and Category.

Type of Coherence	Category	Total	Percemt
Relational	1. Implication	2,264	33.91
	2. Adjacency	1,938	29.02
Additive	1. Elaboration	1,426	21.36
	2. Extension	663	9.93
	3. Circumstantial	252	3.78
	4. Metadiscursive	134	2
Total		6,677	100

Table 4. Frequency of Coherence Relations.

Category	Coherence Relations	Frequency	Percentage
Relational/adjacency	Conjunction	1,301	19.48
Additive/elaboration	Description of concept	865	12.95
Relational/adjacency	Sequence	458	6.86
Relational/implication	Cause-effect	354	5.30
Additive/elaboration	Concept specification	305	4.57
Relational/implication	Condition-event	271	4.06
Relational/implication	Claim-argument	269	4.03
Relational/implication	Basic contrast	219	3.28
Relational/implication	Act-purpose	199	2.98
Additive/extension	Specification	198	2.97
Additive/extension	Exemplification	191	2.86
Additive/circumstantial	Mode	181	2.71
Relational/implication	Argument-claim	177	2.65
Additive/metadiscursive	Deictic	134	2.01
Additive/extension	Description	132	1.98
Relational/implication	Purpose-act	99	1.48
Relational/implication	Event-condition	94	1.41
Relational/implication	Instruction-purpose	92	1.38
Additive/elaboration	Concept exemplification	88	1.32
Relational/implication	Purpose-instruction	80	1.21
Additive/extension	Reformulation	81	1.20
Relational/implication	Reason-action	78	1.17
Relational/implication	Claim-question	72	1.08
Additive/elaboration	Concept comparison	72	1.08

encyclopedic texts (Redeker & van der Vliet, 2014), contents are organized in a list-like, juxtaposition style. In STs, it is common to find discourse segments that are connected only because they share the same topic, as illustrated in (5) and (6).

(5) En la clase del profesor Fernández hay 11 niños y 16 niñas. El profesor Fernández ha escrito el nombre de cada estudiante en un palillo. (Conjunción, Matemáticas)

'In the class of Professor Fernández, there are 11 boys and 16 girls. Professor Fernandez wrote the name of each student on a stick'. (Conjunction, Mathematics)

(6) El conflicto en una obra literaria es el problema que se plantea. El clímax es el momento más intenso del relato. (Conjunción, Lenguaje)

'In a literary work, the conflict is the problem stated. The climax is the most intense moment of the story'. (Conjunction, Language)

Leaving Conjunction aside and given that the communicative purposes of most KGs analyzed are oriented to teach declarative knowledge (Ibáñez et al., 2017), one would expect that contents are presented to the students via precise definitions and descriptions of concepts, processes, or states. The nature of the other most frequent relations identified seems to be in line with those expectations. Table 4, for instance, shows that among the most frequent types of relations there are two Additives: Concept Description and Concept Specification. Examples (7) and (8) illustrates a case of Concept Description, a relation in which one event provides features of a concept introduced in the other event.

(7) Las leyendas son relatos fantásticos, a veces basados en historias reales, ocurridas en un tiempo no muy lejano. Se transmiten de boca en boca y de generación en generación. (Descripción de Concepto, Lenguaje)

'Legends are fantasy stories, sometimes based on real events, occurring in a not-too-distant time. They are transmitted orally, from generation to generation'. (Concept Description, Language)

(8) Los castores son una de las especies de roedores más grandes del planeta. Se caracterizan por tener una gran habilidad para construir estructuras. (Descripción de Concepto, Ciencias)

'Beavers are one of the largest species of rodents on the planet. They are characterized by their great ability to build structures'. (Concept Description, Science)

In addition, when teaching declarative knowledge of any field, explanations of the hows and whys of what is being described are required. Those explanations often involve causality (Ohlsson, 2002; Salmon, 1998). Table 4 shows that those explanations are provided in STs mainly through specific types of causal relations such as Cause-Effect, a relation in which the occurrence of one event leads to the occurrence of another, and Condition-Event, a relation in which the (non) occurrence of an event may lead to the (non) occurrence of another. Examples (9) and (10) illustrate a Cause-Effect and Condition-Event relation, respectively.

(9) Durante la noche, los estomas se cierran producto de la disminución en la concentración de glucosa; por ende, la transpiración disminuye en relación con el día. (Causa-Efecto, Ciencias)

'During the night, stomata close due to the decrease in glucose concentration; therefore, transpiration decreases in relation to the time of day'. (Cause-Effect, Science)

(10) Si se duplica la altura del cono, también se duplica el volumen. (Condición-Evento, Matemáticas)

'If the height of the cone is doubled, the volume is also doubled'. (Condition-Event, Mathematics)

As already mentioned, the communicative purpose of STs is oriented towards the instruction of both declarative and procedural knowledge. Indeed, one of the KGs analyzed (Procedural Guide) aims at instructing about the execution of procedures of a given discipline. Therefore, it seems reasonable to expect that contents are sequenced to teach students the instructions and steps required to carry out a procedure or solve a problem (Özsoy, Kuruyder, & Çakiroglu, 2015). Table 4 shows that those steps or instructions are exposed through Sequence, the third highest frequency, as illustrated in Examples (11) and (12).

(11) Primero, identifica las características de los datos de cada grupo. Luego, establece criterios de comparación. (Secuencia, Matemáticas)

'First, identify the characteristics of the data of each group. Then, establish the comparison criteria'. (Sequence, Mathematics)

(12) Escurran el agua y luego, dejen enfriar el huevo una vez que esté cocido. (Secuencia, Ciencias)

'Drain the water and then, allow the egg to cool down once it is boiled'. (Sequence, Science)

Based on these results, one might infer that these are the six prototypical types of coherence relations in STs and, consequently, claim they are the ones through which knowledge is predominantly communicated in primary STs. However, and based on the assumption of disciplinary variation (Bhatia, 2004; Hyland, 2000, 2008), we expected the use of coherence relations to vary across school subjects. The next section complements the analysis by describing the distribution of the types of coherence relations across subjects.

Distribution of coherence relations across school subjects

Descriptive statistics were used to determine the frequency of category of coherence relations and how often each appears across subjects. Figure 1 shows the distribution of Relational (Adjacency and Implication) and Additive Coherence (Elaboration, Extension, Circumstantial and Metadiscursive) relations across subjects.

After visually identifying potential variation, a two-way independence χ^2 test was performed to assess the relationship between category of coherence relation and school subject. The results were statistically significant showing a relationship between category and the school subject (χ^2 (15) = 345,101, p = <.01) with an effect size of w = .226, which can be interpreted as near moderate (Cohen, 1988). To find the individual differences, Bonferroni correction was used resulting in an

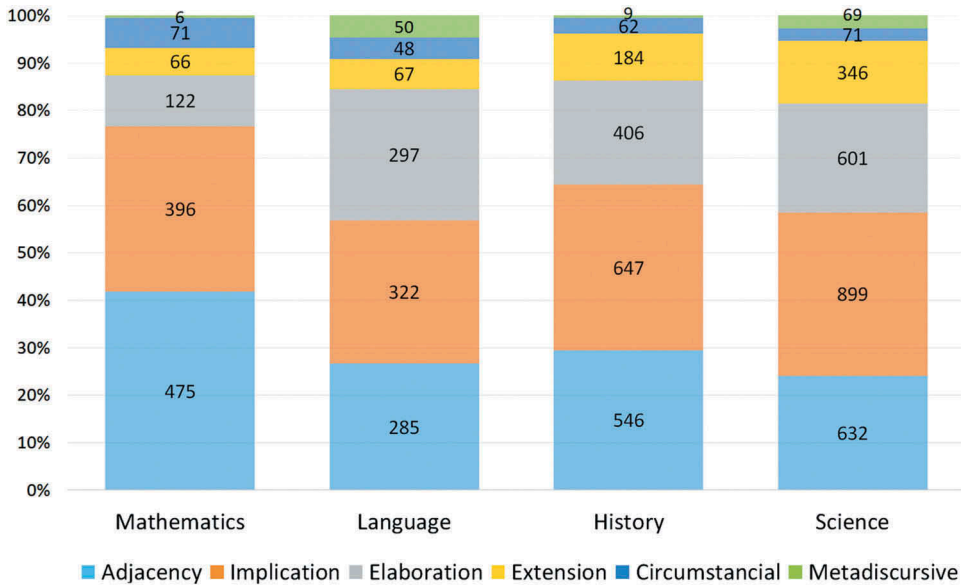


Figure 1. Distribution of categories across school subjects.

adjusted p value of .002 (MacDonald & Gardner, 2000). Based on this value, all pairwise comparisons were tested between school subjects and categories of coherence.

In History, Metadiscursive coherence relations have fewer occurrences than expected ($z = -5.50$, $p < .002$). On the contrary, in Language this category ($z = 6.79$, $p < .002$) and Elaboration ($z = 5.59$, $p < .002$) seem to be important as they appear more than expected. Language does not seem to be particularly related to Extension, since its occurrence is less than expected ($z = -4.37$, $p < .002$). Something different occurs in Science, where Extension shows more than expected ($z = 7.21$, $p < .002$). Two categories appear less than expected in Science. These are Adjacency ($z = -7.06$, $p < .002$) and Circumstantial ($z = -3.66$, $p < .002$). Finally, Mathematics seems to be the most specifically configured school subject, considering it shows association with five of the six categories. Elaboration ($z = -9.59$, $p < .002$), Extension ($z = -5{,}10$, $p < .002$), and Metadiscursive ($z = -3.90$, $p < .002$) appear less than expected, whereas Adjacency ($z = 10.42$, $p < .002$) and Circumstantial ($z = 4.8$, $p < 0.002$) show more than expected.

Although the χ^2 test of independence has shown significant relations between some categories and school subjects, this analysis does not allow us to observe the possible relation between types of coherence relations and school subjects. The relation between each school subject and the types of coherence relations was also analyzed. Table 5 shows the six most frequent types of coherence relations per subject.

Table 5. Most Frequent (Prototypical) Coherence Relations per Subject.

	Science			History			Language			Mathematics	
	N	%		N	%		N	%		N	%
Conjunction	440	16.8	Conjunction	410	22.11	Conjunction	215	20.11	Conjunction	236	20.77
Conc. descr.	341	13.03	Conc. descr.	283	15.26	Conc. descr.	154	14.41	Sequence	220	19.37
Cause-effect	184	7.03	Cause-effect	152	8.19	Conc. specific	79	7.39	Conc. descr.	87	7.66
Cond.-event	172	6.57	Claim-argument	104	5.52	Act-purpose	60	5.61	Mode	70	6.16
Conc. specif	144	5.5	Basic contrast	93	5	Claim-argument	59	5.5	Instr.-purpose	67	5.89
Sequence	129	4.93	Act-purpose	84	4.53	Deictic	50	4.68	Cond.-question	54	4.75

As shown in Table 5 some types of coherence relations are common to all subjects, whereas others are specific to only some of them. Conjunction is the most frequent type of coherence relation across subjects, which is in accordance with its predominance in the corpus. These data and previous studies (Ibáñez & Moncada, 2017; Redeker & van der Vliet, 2014) support the idea that regardless of the type of discourse, genre or register, Conjunction is fundamental in discourse organization.

Table 5 also shows that Concept Description is common to all subjects occurring with a high frequency. In fact, in Language, History, and Science this relation has the second highest frequency, whereas in Mathematics it is the third. Hence, the findings suggest that additive relations seem to play a prominent role in pedagogic discourse, more specifically in KGs, as they allow the details about a concept to be further explored. Examples (13), (14), (15), and (16) illustrate cases of Concept Description in Science, History, Language, and Mathematics, respectively.

(13) El oxígeno es esencial para la vida. Su disminución brusca genera importantes alteraciones que pueden llevar incluso a la muerte. (Descripción de Concepto, Ciencias)

'Oxygen is essential for life. Its abrupt decrease generates important alterations that can even cause death.' (Concept Description, Science)

(14) La Zona Central se caracteriza por desarrollar actividades económicas vinculadas a los servicios y la industria. Además, es centro de las actividades culturales, científicas y educacionales del país. (Descripción de Concepto, Historia)

'The Central Zone is characterized by developing economic activities that are linked to services and industry. Additionally, it is the center of the cultural, scientific and educational activities for the whole country.' (Concept Description, History)

(15) Su estructura tradicional es de grupos de versos de ocho sílabas, cuyos versos pares poseen rima asonante. Una de sus características principales es que narra una historia. (Descripción de Concepto, Lenguaje)

'Its traditional structure are groups of eight syllables verses, where even verses have assonance rhyming. One of its main characteristics is that it tells a story.' (Concept Description, Language).

(16) La marca de clase es el valor medio de cada intervalo. Equivale a la semisuma entre el límite superior y el límite inferior de cada intervalo. (Descripción de Concepto, Matemáticas)

'The class interval is the average value of each interval. It is equal to the half of the sum between the upper limit and the lower limit of each interval.' (Concept Description, Mathematics)

Furthermore, it can be noted that Science and History, despite apparently coming from different knowledge domains (Hyland, 2000), share the types of coherence relations of Conjunction, Concept Description, and Cause-Effect as the first three most frequent. A plausible explanation for this situation could be related to the role objectivity plays in educational contexts, particularly in Science (Davson-Galle, 2002; Gauld, 1982; Schwartz & Lederman, 2008) and History (Cleophas, 2014, Cruse, 2011; Hammarlund, 2012). As illustrated in (17) and (18), both disciplines seek to disseminate their knowledge mainly through objective descriptions of events, minimizing subjective appreciations.

(17) En un líquido, las moléculas se atraen mutuamente, y cada una de ellas está rodeada por moléculas iguales. Por ello, la fuerza de atracción será la misma en todas direcciones. (Causa-Efecto, Ciencias)

'In a liquid, molecules attract each other, and each one is surrounded by the exact same molecules. Therefore, the force of attraction will be the same in all directions.' (Cause-Effect, Science)

(18) En 1455, Portugal obtuvo del papa Nicolás V la propiedad de todas las tierras descubiertas "desde cabo Bojador hacia Guinea y más allá" (Doc. 4), lo que provocó conflictos con los reinos de Castilla y Aragón. (Causa-Efecto, Historia)

'In 1455, Portugal obtained the ownership of all the lands discovered "from Cape Bojador to Guinea and beyond" from Pope Nicolás V (Doc. 4), a fact that caused conflicts with the kingdoms of Castile and Aragon.' (Cause-Effect, History)

With regard to the types of coherence relations that characterize certain school subjects, the corpus analysis has revealed interesting patterns. Sequence, for example, has a low frequency in History (4.1%), Language (3.09%), and Science (4.93%), but in Mathematics is the second most frequent relation (19.37%). The predominance of this relation in Mathematics is in accordance with the procedures students are expected to learn and master in that subject. In fact, following the appropriate steps to solve problems is essential to mathematics (Özsoy et al., 2015). Examples (19) and (20) illustrate this finding:

(19) Marca las intersecciones de las rectas perpendiculares con el eje de reflexión. Luego con el compás, desde el punto de intersección entre el eje de simetría y la perpendicular, mide la distancia hasta el punto original. (Secuencia, Matemáticas)

'Mark the intersections of the perpendicular lines with the reflection axis. Then, with the compass, from the point of intersection between the axis of symmetry and the perpendicular line, measure the distance to the original point.' (Sequence, Mathematics)

(20) Escribe cada término de la expresión algebraica como el producto de sus factores. Luego, identifica el factor o grupo de factores que es común a todas las descomposiciones. (Secuencia, Matemáticas)

'Write each term of the algebraic expression as the product of its factors. Then, identify the factor or group of factors common to all breaking downs'. (Sequence, Mathematics)

Mode, Instruction-Purpose, and Condition-Question are also prototypical of Mathematics, which are the fourth, fifth, and sixth most commonly occurring relations in that subject, respectively. The frequency of these relations could be explained by the role they play in the development of mathematical reasoning. In this subject matter, students are frequently provided with instructions regarding how to do a calculation (Mode), as in (21), and the reason why an action or step of a process should be carried out (Instruction-Purpose), as in (22). Besides, to verify students' learning, some problem-solving activities present hypothetical situations followed by a question, whose solution requires the students apply what they have learnt through the lesson or unit (Condition-Question), as illustrated in (23).

(21) Calcula el producto usando el algoritmo conocido para multiplicación de números decimales, considerándolos como valores positivos. (Modo, Matemáticas)

'Calculate the product using the algorithm known for decimal numbers multiplication.' (Mode, Mathematics).

(22) Expresa la relación como porcentaje para realizar una comparación de cantidades considerando un total de 100. (Instrucción-Propósito, Matemáticas)

'Express the relation as a percentage to perform a comparison of quantities considering a total of 100.' (Instruction-Purpose, Mathematics)

(23) Si el presidente y alcalde asumen el mismo año ¿en cuántos años más podrán presentarse a la elección juntos? (Condición-Pregunta, Matemáticas)

'If the president and mayor take office the same year, how many more years will they be able to stand for election together again?' (Condition-Question, Mathematics)

The emphasis on coherence relations that involves straightforward instructions seem to be worrisome considering the shift from giving direct steps to students toward the implementation of creative open-ended tasks in mathematics that has taken place in the last decade (Ayllón, Gómez, & Ballesta-Claver, 2016; Mueller, Yankelewitz, & Maher, 2014; Norqvist, 2016).

Condition-Event is another case of a coherence relation type which is specific to a subject. As shown in Table 5, it is the fourth most frequently occurring type of relation in Science. Given that most contents covered in this subject describe different natural systems and their processes (Hoffman & Nitecki, 1987), it seems reasonable to find explanations regarding how some biological

processes are triggered when certain circumstances are (not) met. Examples (23) and (24) show typical cases of Condition-Event in Science.

(24) Si esto se produce, los oviductos conducen el embrión (etapa inicial del desarrollo) hasta el útero. (Condición-Evento, Ciencias)

'If this happens, oviducts lead the embryo (initial stage of development) to the uterus.' (Condition-Event, Science)

(25) Si el ovocito no es fecundado, el cuerpo lúteo degenera. (Condición-Evento, Ciencias)

'If the oocyte is not fertilized, the corpus luteum degenerates'. (Condition-Event, Science)

Overall, these findings suggest that although coherence relations are part of pedagogic discourse in general, their frequency patterns vary depending on the subject. Mathematics showed the most specific configuration as it presents a high frequency of three specific coherence relations that are not frequent in the other subjects examined. This is consistent with the results of the statistical χ^2 tests that indicated a statistically significant relation between categories and subjects. In addition, the specificity of mathematical discourse is in agreement with studies that have analyzed mathematics as its own form of discourse from other lines of research (Gorgorió & Planas, 2001; Güçler, Wang, & Kim, 2015; O'Halloran, 2004). Collectively, our data from primary STs support the claim that academic disciplines have different conventional ways of communication (Flowerdew, 2002; Hyland, 2008) and that these differences are reflected in the frequency in which discursive phenomena occur across disciplines. Finally, bearing in mind that coherence relations differ in their internal complexity, our results suggest there are school subjects in which students have to continuously face types of coherence relations that have proved to be difficult. In History, for instance, students have to deal with relations that involve a high degree of subjectivity (Claim-Argument) and negative polarity (Basic Contrast). Therefore, we could assume that in this subject students' deep comprehension and learning from text might be hindered.

Conclusions

Driven by the assumption that to learn from STs, students need to successfully process those coherence relations they find in their STs (van Silfhout et al., 2015), the present study addressed two main questions. First, is knowledge communicated through particular types of coherence relations in STs? Second, do the types of coherence relations used to communicate knowledge vary depending on school subject?

The manual analysis carried out allowed us to obtain information related to both questions. Regarding the first one, we identified the categories and types of coherence relations used in STs of four primary school subjects. Results showed the existence of 40 types of coherence relations which were grouped into six categories (Adjacency, Implication, Elaboration, Extension, Metadiscursive, and Circumstantial). Moreover, 53% of the total number of instances of coherence relations identified was concentrated only in six types (Conjunction, Concept Description, Sequence, Concept Specification, Cause-Effect, and Condition-Event). This evidence allowed us to believe that these are the types of coherence relations through which knowledge is communicated in the STs we analyzed.

Regarding the second question, we could determine that while there were certain types of relations that were frequently used across the four school subjects (Conjunction and Concept Description), others showed to be specific (Condition-Event in Science, Basic Contrast in History, Deictic in Language, and Condition-Question in Mathematics). We also observed variation among the most frequent types per subject. For instance, while Sequence is the second most frequent relation in Mathematics (19.37%) and the sixth in Science (4.93%), they are not even ranked among the top six in Language and History. Further examples are Condition-Event and Mode, the fourth

most frequent types of relation in Science (6.57%) and Mathematics (6.16%), respectively, but rather low in the other subjects.

The use of a top down-bottom up framework also allowed us to identify types of relations, which on the one hand are prototypical of pedagogic discourse and on the other had not been described yet (Claim-Question, Condition-Instruction, Condition-Question, Reason-Instruction, and Instruction-Purpose). These findings not only prove that coherence relations inventories are open, but also confirm that by using this type of approach, taxonomies can be extended and adapted to the corpus being analyzed, resulting in more comprehensive tools. In the case of our study, two of the new types of relations identified occur most frequently in Mathematics (Instruction-Purpose (5.89%) and Condition-Question (4.75%). Further research is needed to determine whether this top down-bottom up framework allows the identification of new coherence relations when describing other genres used in different contexts.

However, these results should be interpreted keeping possible limitations in mind. First, we did not include texts belonging to Curriculum Genres in our corpus. Although they are not the genres used to communicate knowledge, they are part of the learning process in classroom instruction. Hence, to account for the use of coherence relations in primary STs, we might have included them. Accordingly, further research, in which we expand the corpus by adding these discourse genres, is needed to show whether the results we have obtained are confirmed. Another limitation addresses the strategies we adopted during the annotation process. Coders were trained for 4 months and they performed double coding, but the intercoder reliability obtained was not as high as expected. To different specialists, annotation has proven to be a difficult task, which is reflected in low intercoder agreement scores (Poesio & Artstein, 2005). For that reason, to improve intercoder reliability, innovative strategies have been implemented, which include double coding, one-coder-does-all, and even, working with nontrained, nonexpert coders (Spooren & Degand, 2010; Scholman et al., 2016). In future research, we must explore some of these new annotation strategies.

While it is true that some aspects of our study could be improved in future research, the identification of the most frequent coherence relations in STs and their distribution across school subjects can be understood as a contribution. Following the cumulative cognitive complexity approach (Evers-Vermeul & Sanders, 2009; Spooren & Sanders, 2008), these findings provide relevant information about the relation between school subjects and degree of complexity of coherence relations. Our data showed that while in some subjects the most frequent types of coherence relations are those considered difficult (Claim-Argument, Basic Contrast, and Act-Purpose, in History), in other subjects the most frequent types of relations are those assumed to be less difficult (Cause-Effect, Condition-Event, Concept Specification, and Sequence, in Science). If we bear in mind not only the central role that establishing coherence relations plays in comprehension but also that recent studies have proved that primary school children exhibit difficulties in processing some types of coherence relations (Knoepke et al., 2017), our findings have important implications for classroom instruction, STs design, and text comprehension research in Chile. According to our results, there is a strong relation between school subject and type of coherence relation. Hence, the teaching-learning process could be facilitated by providing students with reading strategies, ad-hoc to the types of coherence relations they more frequently find in different school subjects. This approach may be suitable to aid primary school students to learn from written texts across disciplines. Regarding text comprehension research, it seems necessary to carry out empirical studies in order to answer questions such as: how are the most frequent relations in different school subjects processed by Chilean primary school students? Are the most frequent relations processed differently from those that are less frequent?

Taken together, our findings represent interesting insights into the way coherence relations are used in Primary School textbooks in Chile. Further studies are now needed to investigate whether similar patterns can also be observed in high STs and in different languages. This would provide further information on the process of knowledge recontextualization across different educational levels (Bernstein, 1990).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by grant FONDECYT 1160094 from the National Commission for Scientific and Technological Research (CONICYT).

ORCID

Benjamín Cárcamo http://orcid.org/0000-0001-7330-9007

References

Atienza, E. (2007). Discurso e ideología en los libros de texto de ciencias sociales. *Discurso & Sociedad*, *1*(4), 543–574.

Ayllón, M., Gómez, I., & Ballesta-Claver, J. (2016). Mathematical thinking and creativity through mathematical problem posing and solving. *Propósitos y representaciones*, *4*(1), 169–218.

Barletta, N., & Chamorro, M. (Eds.). (2011). *El texto escolar y el aprendizaje: Enredos y desenredos*. Barranquilla, Colombia: Universidad del Norte.

Barrios, B. (2012). Actividades de lectura y escritura en el texto escolar El Cardenalito para primer grado. *Investigación y postgrado*, *27*(1), 111–138.

Bernstein, B. (1990). *The structuring of pedagogic discourse: Class, codes & control* (Vol.IV). London: Routledge.

Bhatia, V. (2004). *Worlds of written discourse: A genre based view*. London, UK: Continuum.

Biber, D., Connor, U., & Upton, T. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. Amsterdam, Netherlands: John Benjamins.

Black, J., & Bower, G. H. (1980). Story understanding as problem-solving. *Poetics*, *9*, 223–250. doi:10.1016/0304-422X(80)90021-2

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, *7*(2), 235–261.

Bourdieu, P. (1999). *Contrafuegos. Reflexiones para servir a la resistencia contra la invasión neoliberal*. Barcelona, Spain: Anagrama.

Bublitz, W. (1998). Cohesion and coherence. In J. Verschueren, J. Östman, J. Blommaert, & C. Bulcaen (Eds.), *Handbook of pragmatics* (pp. 1–15). Amsterdam, Netherlands: John Benjamins.

Cain, K., Patson, N., & Andrews, L. (2005). Age-and ability-related differences in young readers' use of conjunctions. *Journal of Child Language*, *32*(4), 877–892.

Canestrelli, A., Mak, W., & Sanders, T. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes*, *9*, 1394–1413. doi:10.1080/01690965.2012.685885

Chamorro, D., Mizuno, J., & Moss, G. (2011). Tergiversaciones y correspondencias: La metáfora y sus bemoles. In D. Chamorro & N. Barletta (Eds.), *El texto escolar y el aprendizaje. Enredos y desenredos* (pp. 101–122). Barranquilla, Colombia: Editorial Universidad del Norte.

Christie, F. (2002). *Classroom discourse analysis*. London, UK: Continuum.

Cleophas, F. (2014). Writing and contextualising local history. A history narrative of the Wellington horticultural society (Coloured)'. *Yesterday and Today*, *11*, 21–53.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological. *Measurement*, *20*, 37–46.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *TESOL Quarterly*, *42*, 475–493. doi:10.1002/tesq.2008.42.issue-3

Cruse, I. (2011). *Debate on 20 October: Teaching of history in schools, house of lords library note*. Retrieved December 09, 2018, from https://www.parliament.uk/documents/lords-library/Library%20Notes/2011/LLN%202011-030%20TeachingHistorySchoolsFP2.pdf

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, *20*(6), 445–464. doi:10.1007/BF01067637

Das, D., & Taboada, M. (2017). Signalling of coherence relations in discourse, beyond discourse markers. *Discourse Processes*, *55*(8), 1–29.

Davson-Galle, P. (2002). Science, values and objectivity. *Science and Education*, *11*(2), 191–202. doi:10.1023/A:1014412500803

Dufty, D., Graesser, A., Louwerse, M., & McNamara, D. (2006). Assigning grade levels to textbooks: Is it just readability? In R. Son (Ed.), *Proceedings of the 28th annual meetings of the cognitive science society* (pp. 1251–1256). Mahwah, NJ: Erlbaum.

Evers-Vermeul, J. (2005). *The development of Dutch connectives: Change and acquisition as windows on form-function relations* (Unpublished PhD dissertation). Utrecht University, LOT, Utrecht, The Netherlands.

Evers-Vermeul, J., & Sanders, T. (2009). The emergence of Dutch connectives; how cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language*, *36*(4), 829–854. doi:10.1017/S0305000908009227

Evers-Vermeul, J., & Sanders, T. (2011). Discovering domains – On the acquisition of causal connectives. *Journal of Pragmatics*, *43*(6), 1645–1662. doi:10.1016/j.pragma.2010.11.015

Fleiss, J. (1981). *Statistical methods for rates and proportions*. London, UK: John Wiley & Sons.

Flowerdew, J. (2002). *Academic discourse*. London and New York, NY: Routledge.

Gauld, C. (1982). The scientific attitude and science education: A critical reappraisal. *Science Education*, *66*, 109–121. doi:10.1002/(ISSN)1098-237X

Gorgorió, N., & Planas, N. (2001). Teaching mathematics in multilingual classrooms. *Educational Studies in Mathematics*, *47*(1), 7–33. doi:10.1023/A:1017980828943

Graesser, A., & Clark, L. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82–98). New York, NY: Guilford.

Grosz, B., & Sidner, C. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, *12*(3), 175–204.

Güçler, B., Wang, S., & Kim, D. (2015). Conceptualizing mathematics as discourse in different educational settings. *International Education Studies*, *8*(12), 25–32. doi:10.5539/ies.v8n12p25

Hammarlund, K. G. (2012). Promoting procedural knowledge in history education. In D. Ludvigsson (Ed.), *Enhancing student learning in history: Perspectives on university history teaching* (pp. 117–130). Uppsala, Sweden: Uppsala University.

Hoek, J., Evers-Vermeul, J., & Sanders, T. (2017). Segmenting discourse: Incorporating interpretation into segmentation? *Corpus Linguistics and Linguistic Theory*. Ahead of Print Retrieved December 09, 2018, from. doi: 10.1515/cllt-2016-0042

Hoffman, A., & Nitecki, M. H. (1987). Introduction: Neutral models as a biological research strategy. In M. H. Nitecki & A. Hoffman (Eds.), *Neutral models in biology* (pp. 3–8). New York, NY: Oxford University Press.

Hovy, E., Lavid, J., Maier, E., Mittal, V., & Paris, C. (1992). Employing knowledge resources in a new text planning architecture. In R. Dale, E. Hovy, D. Rösner, & O. Stock (Eds.), *Aspects of automated natural language generation* (pp. 57–72). New York, NY: Springer.

Hyland, K. (2000). *Disciplinary discourse: Social interactions in academic writing*. London, UK: Longman.

Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching*, *41*, 543–562. doi:10.1017/S0261444808005235

Ibáñez, R., & Moncada, F. (2017). El resumen de artículos de investigación científica: Variación disciplinar a nivel local y global. *Spanish in Context*, *14*, 273–308. doi:10.1075/sic

Ibáñez, R., Moncada, F., Cornejo, F., & Arriaza, V. (2017). Los géneros del conocimiento en textos escolares de educación primaria. *Calidoscópio*, *15*(3), 462–476. doi:10.4013/cld.2017.153.06

Ibáñez, R., Moncada, F., & Santana, S. (2015). Variación disciplinar en el discurso académico de la biología y del derecho: Un estudio a partir de las relaciones de coherencia. *Onomázein*, *32*, 101–131. doi:10.7764/onomazein.32.6

Knoepke, J., Richter, T., Isberner, M. B., Naumann, J., Neeb, Y., & Weinert, S. (2017). Processing of positive-causal and negative-causal coherence relations in primary school children and adults: A test of the cumulative cognitive complexity approach in German. *Journal of Child Language*, *44*(2), 297–328. doi:10.1017/S0305000915000872

Landis, R., & Koch, G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, *33*, 363–374. doi:10.2307/2529786

Louwerse, M. (2001). An analytic and cognitive parametrization of coherence relations. *Cognitive Linguistics*, *12*(3), 291–315.

MacDonald, P., & Gardner, R. (2000). Type I error rate comparisons of post hoc procedures for I j Chi-square tables. *Educational and Psychological Measurement*, *60*, 735–754. doi:10.1177/00131640021970871

Mann, W., & Thompson, S. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, *8*(3), 243–281. doi:10.1515/text.1.1988.8.3.243

Martin, J., & Rose, D. (2008). *Genre relations: Mapping culture*. London, UK: Equinox.

Martin, J., & Rose, D. (2013). Pedagogic discourse: Contexts of schooling. *RASK: International Journal of Language and Communication*, *38*, 219–264.

Marzábal, A. (2012). Las actividades de los libros de texto de química para la teoría corpuscular y su contribución a la evolución de modelos explicativos. *Estudios Pedagógicos*, *XXXVIII*(1), 181–196. doi:10.4067/S0718-07052012000100011

Maury, P., & Teisserenc, A. (2005). The role of connectives in science text comprehension and memory. *Language and Cognitive Processes*, *20*, 489–512. doi:10.1080/01690960444000151

McNamara, D. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, *55*, 51–62. doi:10.1037/h0087352

McNamara, D., Kintsch, E., Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition Instruction*, *14*, 1–43. doi:10.1207/s1532690xci1401_1

Millis, K., & Just, M. (1994). The influence of connectives on sentence comprehension. *Journal of Memory & Language*, *33*, 128–147. doi:10.1006/jmla.1994.1007

MINEDUC. (2012). *Bases Curriculares para la Enseñanza Básica*. Retrieved December 09, 2018, from http://www.curriculumnacional.cl/inicio/

Mueller, M., Yankelewitz, D., & Maher, C. (2014). Teachers promoting student mathematical reasoning. *Investigations in Mathematics Learning*, *7*(2), 1–20. doi:10.1080/24727466.2014.11790339

Norqvist, M. (2016). *On mathematical reasoning—Being told or finding out* (Unpublished PhD thesis). Umea University, Sweden.

O'Halloran, K. (2004). *Mathematical discourse: Language, symbolism, and visual images*. London, UK: Continuum.

Ohlsson, S. (2002). Generating and understanding qualitative explanations. In J. Otero, J. León, & A. Graesser (Eds.), *The psychology of scientific text comprehension* (pp. 91–128). London, UK: Lawrance Erlbaum Associates.

Otero, J., León, J., & Graesser, A. C. (2002). *The psychology of science text comprehension*. Mahwah, NJ: Erlbaum.

Özsoy, G., Kuruyder, H., & Çakiroglu, A. (2015). Evaluation of student's mathematical problem solving skills in relation to their reading levels. *International Electronic Journal of Elementary Education*, *8*(1), 113–132.

Petrucci, A. (2001). Leer por leer: Un porvenir para la lectura. In G. Cavallo & R. Chartier (Eds.), *Historia de la lectura en el mundo occidental* (pp. 591–625). Madrid, Spain: Taurus.

Poesio, M., & Artstein, R. (2005). The reliability of anaphoric annotation, reconsidered: Taking ambiguity into account. In *Proceedings of the Workshop on Frontiers in Corpus Annotations II: Pie in the sky* (pp. 76–83). Ann Arbor, MI: ACL.

Pretorius, E. (1994). A text linguistic perspective on causality in discourse: Toward a taxonomy of causal relations. *South African Journal of Linguistics*, *12*(sup22), 81–120. doi:10.1080/10118063.1994.9724513

Ramírez, T. (2012). El texto escolar como arma política. Venezuela y su gente: Ciencias sociales 6to. grado. *Investigación y Postgrado*, *27*(1), 163–194.

Redeker, G. (2000). Coherence and structure in text and discourse. In H. Bunt & W. Black (Eds.), *Abduction, belief and context in dialogue. Studies in computational pragmatics* (pp. 233–263). Amsterdam, The Netherlands: John Benjamins.

Redeker, G., & van der Vliet, N. (2014). Explicit and implicit coherence relations in Dutch texts. In H. Gruber & G. Redeker (Eds.), *The pragmatics of discourse coherence: Theory and applications* (Pragmatics & Beyond New Series; Vol. 254, pp. 23–52). Amsterdam, The Netherlands: John Benjamins.

Renkema, J. (2009). *The texture of discourse: Towards an outline of connectivity theory*. Amsterdam, The Netherlands: John Benjamins.

Rose, D. (2014). Analysing pedagogic discourse: An approach from genre and register. *Functional Linguistics*, *1*(11), 1–32.

Sáiz, J. (2011). Actividades de libros de texto de Historia, competencias básicas y destrezas cognitivas, una difícil relación: Análisis de manuales de 1º y 2º de ESO. *Didáctica de las ciencias experimentales y sociales*, *25*, 37–64.

Salmon, W. (1998). *Causality and explanation*. New York, NY: Oxford University Press.

Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes*, *24*, 119–148. doi:10.1080/01638539709545009

Sanders, T., & Noordman, L. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *29*, 37–60. doi:10.1207/S15326950dp2901_3

Sanders, T., & Pander Maat, H. (2006). Cohesion and coherence. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (Vol. 2, pp. 591–595). Amsterdam, The Netherlands: Elsevier.

Sanders, T., Spooren, W., & Noordman, L. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15* (1), 1–35. doi:10.1080/01638539209544800

Sanders, T., Spooren, W., & Noordman, L. (1993). Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*, *4*(2), 93–133. doi:10.1515/cogl.1993.4.2.93

Scholman, M., Evers-Vermeul, J., & Sanders, T. (2016). A step-wise approach to discourse annotation: Towards a reliable categorization of coherence relations. *Dialogue & Discourse*, *7*(2), 1–28.

Schwartz, R., & Lederman, N. (2008). What scientists say: Scientists' views of nature of science and relation to science context. *International Journal of Science Education*, *30*(6), 727–771. doi:10.1080/09500690701225801

Spooren, W., & Degand, L. (2010). Coding coherence relations: Reliability and validity. *Corpus Linguistics and Linguistic Theory*, *6*(2), 241–266. doi:10.1515/cllt.2010.009

Spooren, W., & Sanders, T. (2008). The acquisition of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, *40*(12), 2003–2026. doi:10.1016/j.pragma.2008.04.021

Taboada, M., & Das, D. (2013). Annotation upon annotation: Adding signalling information to a corpus of discourse relations. *Dialogue and Discourse*, *4*(2), 249–281. doi:10.5087/dad.2013.211

Taboada, M., & Gómez-González, M. (2012). Discourse markers and coherence relations: Comparison across markers, languages and modalities. *Linguistics and the Human Sciences*, *6*(1–3), 17–41. doi:10.1558/lhs.v6i1-3.17

Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, *24*, 595–611. doi:10.1016/0749-596X(85)90048-8

Traxler, M., Bybee, M., & Pickering, M. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology*, *50*, 481–497. doi:10.1080/027249897391982

van Den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423–446). Hillsdale, NJ: Erlbaum.

van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2015). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes*, *52*(1), 47–76. doi:10.1080/0163853X.2014.905237

Zufferey, S., & Gygax, P. M. (2016). The role of perspective shifts for processing and translating discourse relations. *Discourse Processes*, *53*(7), 532–555. doi:10.1080/0163853X.2015.1062839

Appendix A

Categories and Types of Coherence Relations

Additive Coherence			
Elaboration	Extension	Circumstantial	Metadicursive
Concept definition	Exemplification Reformulation	Time	Deictic
Concept example	Specification	Simultaneity	
Concept specification	Comparison	Mode	
Concept reformulation	Description		
Concept comparison			
Concept description			

Relational Coherence	
Adjacency	Implication
Conjunction	Cause-effect
Sequence	Effect-cause
Opposition	Basic contrast
Disjunction	Nonbasic contrast
Substitution	Event-condition
	Condition-event
	Reason-action
	Action-reason
	Act-purpose
	Purpose-act
	Argument-claim
	Claim-argument
	Evidence-deduction
	Instruction-purpose
	Purpose-instruction
	Claim-question
	Condition-question
	Condition-instruction
	Instruction-reason
	Reason-instruction

Appendix B

Sample of the list of School Textbooks

Subject	Year	Grade	Publishing company
LANGUAGE	2012	5	Cal y Canto
		6	Zig-Zag S.A.
		7	Mn Editorial Ltda.
	2013	6	Editorial Norma de Chile S.A.
		7	Santillana
		8	Santillana
	2015	7	Santillana
		8	Santillana
	2016	5	Piedra de Sol
		6	Piedra de Sol
		7	Santillana
		8	Ediciones SM
MATHEMATICS	2012	5	Santillana
		6	Mn Editorial Ltda.
		7	Santillana
	2013	8	Santillana
	2014	5	Galileo Editorial
		6	Galileo Editorial
		8	Galileo Editorial
	2015	7	Galileo Editorial
		8	Galileo Editorial
	2016	5	Galileo Editorial
		6	Galileo Editorial
		7	Ediciones SM
		8	Ediciones SM

Appendix C

Example of a KG (Procedural Guide Extracted from a Mathematics ST)

Situación Representar hechos cotidianos

Al inicio de la sección, cuando hablamos del *kril*, se destacaron algunas palabras necesarias para completar la información, como bajo cero o bajo el nivel del mar.

¿Cómo podemos representar este tipo de información?

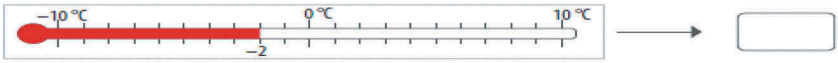
Paso 1 Identifica la información.

La temperatura del agua puede ser de 2 °C **bajo cero**.

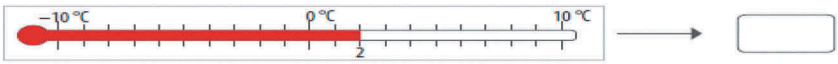
La temperatura del agua también puede ser de 2 °C **sobre cero.**

Paso 2 Asocia a una de esas expresiones un número negativo y a la otra, un número positivo y represéntalas en un termómetro.

La temperatura del agua puede ser de 2 °C **bajo cero**.



La temperatura del agua también puede ser de 2 °C **sobre cero**.



Así, conceptos como bajo el nivel del mar, gastos y déficit se indican con el signo _____. Por el contrario, sobre el nivel del mar, ganancias y superávit se asocian con el signo _____.

Appendix D

Frequency of the 40 Types of Coherence Relations Identified

Type of Coherence/Categories	Coherence Relations	Frequency	Percent
Relational/adjacency	Conjunction	1,301	19.5
Additive/elaboration	Concept description	865	13.0
Relational/adjacency	Sequence	458	6.9
Relational/implication	Cause-effect	354	5.3
Additive/elaboration	Concept specification	305	4.6
Relational/implication	Condition-event	271	4.1
Relational/implication	Claim-argument	269	4.0
Relational/implication	Basic contrast	219	3.3
Relational/implication	Act-purpose	199	3.0
Additive/extension	Specification	198	3.0
Additive/extension	Exemplification	191	2.9
Additive/circumstantial	Mode	181	2.7
Relational/implication	Argument-claim	177	2.7
Additive/metadiscursive	Deictic	134	2.0
Additive/extension	Description	132	2.0
Relational/implication	Purpose-act	99	1.5
Relational/implication	Event-condition	94	1.4
Relational/implication	Instruction-purpose	92	1.4
Additive/elaboration	Concept exemplification	88	1.3
Relational/implication	Purpose-instruction	80	1.2
Additive/extension	Reformulation	81	1.2
Relational/implication	Reason-action	78	1.2
Relational/implication	Claim-question	72	1.1
Additive/elaboration	Concept comparison	72	1.1
Relational/implication	Condition-question	66	1.0
Relational/adjacency	Opposition	65	1.0
Relational/adjacency	Disjunction	63	0.9
Additive/extension	Comparison	62	0.9
Additive/elaboration	Concept definition	61	0.9
Relational/implication	Effect-cause	58	0.9
Relational/adjacency	Substitution	51	0.8
Additive/circumstantial	Temporary	41	0.6
Additive/elaboration	Concept reformulation	34	0. 5
Additive/circumstantial	Simultaneity	30	0.4
Relational/implication	Condition-instruction	28	0.4
Relational/implication	Evidence-deduction	28	0.4
Relational/implication	Action-reason	26	0.4
Relational/implication	Instruction-reason	22	0.3
Relational/implication	Nonbasic contrast	17	0.3
Relational/implication	Reason-instruction	15	0.2
	Total	6,677	100%