The PhD Project Description.

DIDI – Discovering Discourse: The acquisition of discourse connectives in L1

1. State of research: Connectives in first language acquisition

A body of literature has addressed the question of the early acquisition of connectives, and more specifically the order in which different connectives are mastered (e.g. Bloom et al., 1980; Braunwald, 1985; Evers-Vermeul & Sanders, 2009). This order is now rather well validated and is similar across languages. Children start producing additive connectives, followed by temporal ones, then by causal ones, and finally produce adversative connectives. The factors called upon to explain this order of acquisition depend on the underlying theoretical background of the studies, some of them emphasizing syntactic complexity, others insisting on the role of the cognitive complexity of the coherence relation and others stressing the role of parental input. Based on a literature review, Geva (2006) argues that both syntactic and semantic factors influence the acquisition process of connectives, and Evers-Vermeul & Sanders (2009) also advocate a multidimensional explanation. All the studies cited above have investigated the emergence of children's spontaneous production of connectives in naturalistic data. However, spontaneously occurring connectives represent only one aspect of the acquisition process. These studies do not provide insights into the way children understand and manipulate these connectives in specific contexts, nor do they provide information about how the acquisition continues after the early emergence of these connectives in children's speech. Indeed, other studies have revealed that the appropriate use of connectives and their understanding (McClure & Geva, 1983; Irwin & Pulver, 1984) are developing over many years after the connectives are first used. In a set of experiments, Cain & Nash (2011) tested the way 8- and 10-year-old children understand and process frequent connectives such as *because, and, but*, or *after*. Their results indicate that children as late as 10-year-old perform significantly less well than adults in a cloze task assessing their use of temporal, causal and adversative connectives. In another experiment assessing children's evaluation of connective meanings, 8-year-old children performed less well than adults for all connectives and 10- year-olds still had an inferior performance only for temporal connectives. In a set of experiments targeting more specifically the understanding of temporal relations, Pyykkönen and Järvikivi (2012) demonstrated that children aged 12 still experience difficulties understanding temporal relations when the connective used (i.e. *after*) implies a reversed order of the related segments with respect to the chronological order of events in the world. In a nutshell, children still experience many difficulties with connectives at the end of primary school. Their later development during secondary and high-school is, however, understudied in the literature. In particular, the age from which children display an adult-like behavior in the processing and understanding of connectives remains largely to be established.

However, young readers are partially able to use the clues provided by connectives during reading. For instance, at the age of 8, children read two sentences that are related by an appropriate connective faster than the same sentences related with an inappropriate or neutral connective (namely *and*) or not related by a connective at all (Cain & Nash, 2011). The role of connectives for discourse processing also appears to be variable across discourse relations for children. Indeed, they appear to understand connectives encoding positive coherence relations such as causal and additive connectives better compared to connectives encoding negative relations such as concession and contrast in off-line measures of comprehension (Geva, 2006; Knoepke et al., 2017). In sum, even during the school years, cognitive complexity still appears to play a key role for children's understanding of connectives. But this is not the only factor. An alternative explanation is that children specifically fail to understand infrequent connectives, which are unfamiliar lexical items, simply due to vocabulary limitations. At the age of 9 to 10, connective familiarity and cognitive complexity were both found to influence the ability of fourth graders coming from a language minority

background to understand connectives (Crosson et al., 2008).

In the DIDI project, we will test the familiarity hypothesis further by investigating the way native speaking secondary-school and high-school teenagers progressively come to integrate the meaning of a range of connectives restricted to the written mode (see Experiment 1). Previous studies always tested connectives communicating one specific discourse relation. Yet, many connectives are polyfunctional, i.e. they can convey distinct discourse relations depending on context. For example, the connective *since* encodes either a causal or a temporal relation depending on context. Polyfunctionality has been found to be a limiting factor for adult native speakers' ability to judge the appropriateness of connectives, especially when a connective is used in an infrequent function. For example, English native speakers perform at ceiling when they have to evaluate the correct and incorrect uses of *while* as a temporal marker but not as a contrastive marker (Zufferey et al., 2015). DIDI will test the role of polyfunctionality as a factor influencing children's understanding of connectives used in the written mode.

Finally, embedded discourse relations, illustrated in (1), may add to the difficulty of discourse relations, but have never been investigated.

(1) Peter is sad because, if his mother finds out about his grade, he will be punished.

In corpus data, many discourse relations are embedded into one another (Hoek et al., 2017). Processing embedded clauses is demanding for working memory (e.g. Traxler, 2007) and this might impair teenagers' ability to understand coherence relations when embedding is involved. The role of embeddedness has however not been assessed in relation to teenagers' understanding of coherence relations. DIDI will fill this gap by testing teenagers' ability to process and understand different types of embedded discourse.

We will investigate the factors limiting teenagers' ability to understand and process connectives during their secondary and high-school years, focusing on three underexplored areas of complexity. A first set of experiments will assess the roles of word familiarity, cognitive complexity and polyfunctionality as factors influencing teenagers' ability to understand connectives used in the written mode. A second set of experiments will explore teenagers' ability to process and understand various types of implicit coherence relations. A third set of experiments will assess teenagers' processing and understanding of embedded coherence relations.

2. Method

Example of the Experiment 1a: The acquisition of connectives used in the written mode

From secondary school onwards, students encounter complex texts that contain a variety of connectives, most of which are typically used only in writing. The initial experiments will aim at assessing if and when teenagers develop an understanding of these connectives, and which factors best explain their difficulties. In a previous study (Crosson et al., 2008), word frequency and cognitive complexity were both found to be important to account for young teenagers' difficulties with connectives. However, word familiarity was measured based on the proportion of students who know a word in a given age group, and the connectives tested also included connectives frequent in the spoken mode (*but, because, before*, etc.). Thus, it is not surprising that these connectives were found to be easier than the ones only found in writing. Experiment 1a will include only connectives used in the written mode, in order to assess whether word frequency or cognitive complexity plays the most significant role to explain teenagers' difficulties with these connectives.

Participants

120 French-speaking teenagers will be selected from four different age groups: 30 12-year-olds completing their final year of primary school (8th year in the Swiss Harmos system), 30 14-year-olds during secondary school, 30

16-year-olds who have entered high school, and 30 18-year-olds who finish high-school. A group of 30 adults (University students) will also be tested to act as a control group.

Materials

Experiment 1a will include 4 French connectives. These connectives have been selected based on a corpus study performed by the applicant to compare the frequency of connectives from the Lexconn database of French connectives (Roze et al., 2012) across the written and spoken modes. All the connectives chosen for the experiment were found in written corpora from literary and newspaper genres, but not in spoken French corpora. The list includes: *aussi, en effet, néanmoins, cependant.* The task will be a version of the Text Cohesion Task (Crosson et al., 2008) in which students read two discourse segments related by a blank line and have to insert the correct connective among a choice of four possibilities. Distractor connectives will be chosen so that none of them is possible in the discourse context, but they are nonetheless plausible alternatives. The task will include five correct uses for each connective, as illustrated in (2) for the connective *aussi*.

(2) Marie est partie bien en avance,_____elle n'est pas arrivée en retard à son cours.

Choice of connectives: *aussi, cependant, en revanche, en outre*

Procedure and design

Participants will receive a questionnaire on a piece of paper and will be asked to underline the appropriate connective. The sentences will be in randomized order.

Dependent variables and analyses

The frequency of each connective in written corpora will be used to scale them based on their number of occurrences per million words in a corpus study conducted in preparation of this experiment. A scale of cognitive complexity will be devised based on the four primitives identified by Sanders et al. (1992). For each primitive, one of the alternatives was found to be more complex than the other in the literature (see Hoek et al., 2017 for a review). Thus, complexity on each dimension will be computed by attributing 0 to the easier option and 1 the more complex one, resulting in a complexity score for each connective ranging from 0 to 4. Based on these results, two continuous variables will be used in the experiments: (1) degree of cognitive complexity and (2) connective frequency. In the experiment, accuracy of the connective chosen will be recorded. The role of frequency and cognitively complexity on teenagers' ability to select the correct connective will be analyzed using hierarchical regression analyses, following the methodology used in Crosson et al.(2008).

Hypotheses

According to previous studies, cognitive complexity plays a role for teenagers' ability to use connectives in the early teens but at a later age, only word frequency plays a role, when connectives from both the spoken and the written modes are tested together. Given the age groups that we test, we expect that overall cognitive complexity will play a lesser role compared to word frequency. We also expect the effect of cognitive complexity to diminish with age while word frequency should remain the most relevant factor to explain individual variations even among 18-year-olds.

3. Preparatory steps

October 2019 – December 2019: Literature review and preparation of materials for Experiment 1.

January 2019 – March 2020: Pilot test for Experiment 1.

April 2020 – July 2020: Data collection for Experiments 1.

August – Sept. 2020: Data analysis and presentation of first results of Experiment 1 at a conference.

4. References

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, 235-261.

Braunwald, S. R. (1985). The development of connectives. *Journal of Pragmatics* 9, 513-525.

Cain, K. & Nash, H. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology* 103(2), 429-441.

Canestrelli A., Mak, W. & Sanders, T. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye-movements. *Journal of Language and Cognitive Processes* 28 (9), 1394-1413.

Crosson, A., Lesaux, N., & Martiniello, M. (2008). Factors that influence comprehension of connectives among language minority children from Spanish-speaking backgrounds. *Applied Psycholinguistics* 29, 603–625.

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, 829-854.

Geva, E. (2006). Conjunction use in school children's oral language and reading. In R. Horowitz (Ed.) Talking texts: How speech and writing interact in school learning. Hillsdale, NJ: Erlbaum.

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

Hoek J., Zufferey S., Evers-Vermeul J. & Sanders T. (2017). Cognitive complexity and linguistic marking of coherence relations. A parallel corpus study. *Journal of Pragmatics* 121, 113-131.

Irwin, J., & Pulver, C. (1984). Effects of explicitness, clause order, and reversibility on children's comprehension of causal relationships. *Journal of Educational Psychology* 76, 399-407.

Knott, A. & Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18, 35-62.

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

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

McClure, E., & Geva, E. (1983). The development of the cohesive use of adversative conjunctions in discourse. *Discourse Processes* 6, 411-432.

Pyykkönen, P. & Järvikivi, J. (2012) Children and situation models of multiple events. *Developmental Psychology* 48 (2), 521-529.

Roze, C., Danlos, L. & Muller, P. (2012). Lexconn: a French lexicon of discourse connectives. *Discours* 10, 1-15.

Sanders, T., Land, J. & Mulder, G. (2007). Linguistic markers of coherence improve text comprehension in functional contexts. *Information Design Journal* 15, 219-235.

Sanders, T., Spooren, W. & Noordman, L. (1992). Towards a Taxonomy of Coherence Relations. *Discourse Processes* 15, 1-36.

Traxler, M. (2007). Working memory contributions to relative clause attachment processing: a hierarchical linear modelling analysis. *Memory and Cognition* 35(5), 1107-1121.

Zufferey S., Mak W., Degand L. & Sanders T. (2015). Advanced learners' comprehension of discourse connectives: The role of L1 transfer across on-line and off-line tasks. *Second Language Research* 31(3), 389-411.